

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 04:37:24 ; Search time 1684.26 Seconds  
(without alignments)  
7460.491 Million cell updates/sec

Title: US-09-852-261-1  
Perfect score: 517  
Sequence: 1 ggaccggagacgctctgcgg.....tgaaatacacaagtaaacaat 517

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*

28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result			%				ID	Description
	No.	Score	Match	Length	DB	Query		
	1	344.2	66.6	796	14	CB959991	CB959991	AGENCOURT
c	2	331.6	64.1	558	9	AI503976	AI503976	vm43d08.x
c	3	330.6	63.9	673	12	BM984670	BM984670	UI-CF-EC1
c	4	329.8	63.8	623	9	AW146128	AW146128	um37e10.x
c	5	326.6	63.2	575	9	AI248089	AI248089	qh69f05.x
c	6	316.6	61.2	549	9	AI169253	AI169253	EST215088
c	7	315.8	61.1	558	9	AI265629	AI265629	uj04b07.x
c	8	314.8	60.9	498	9	AA542914	AA542914	ni98c10.s
	9	310	60.0	614	14	CD373004	CD373004	UI-R-GR0-
	10	309	59.8	816	9	AI119218	AI119218	ue94h02.y
	11	303.6	58.7	594	10	BF383724	BF383724	602044632
c	12	299.8	58.0	527	9	AA913900	AA913900	ol35g05.s
c	13	289.6	56.0	642	9	AI876493	AI876493	uj59b10.x
c	14	287.4	55.6	499	9	AW495481	AW495481	UI-M-BH3-
c	15	276	53.4	468	9	AI169770	AI169770	EST215669
	16	274.4	53.1	882	9	AI604642	AI604642	vm43d08.y
c	17	268.2	51.9	430	9	AI478804	AI478804	tm52e04.x
c	18	263.2	50.9	653	13	BQ200567	BQ200567	UI-R-DZ1-
	19	258.4	50.0	608	9	AL599807	AL599807	DKFZp3130
c	20	254.6	49.2	486	9	AA993659	AA993659	ot85g11.s
c	21	254.2	49.2	521	9	AW493459	AW493459	UI-M-BH3-
	22	254.2	49.2	559	12	BI715603	BI715603	ic34h10.y
	23	254.2	49.2	602	13	BU590710	BU590710	AGENCOURT
	24	254.2	49.2	621	12	BI221656	BI221656	602936980
	25	254.2	49.2	1658	11	AK081019	AK081019	Mus muscu
	26	254	49.1	356	9	AW297586	AW297586	UI-H-BW0-
c	27	253.2	49.0	595	9	AI573421	AI573421	mo04b11.x
c	28	252.6	48.9	499	12	BI676839	BI676839	ic56a08.x
c	29	252.6	48.9	500	9	AA945553	AA945553	EST201052
c	30	252.6	48.9	525	9	AA963258	AA963258	UI-R-E1-g
	31	251.4	48.6	482	9	AA456717	AA456717	aa13h06.r
c	32	251	48.5	706	9	AI401719	AI401719	th30b10.x
c	33	249.4	48.2	525	9	AI599751	AI599751	EST251454
	34	248.6	48.1	665	9	AA690767	AA690767	vu57d12.r
	35	247.8	47.9	559	12	BI715465	BI715465	ic33b09.y
	36	247.4	47.9	799	9	AI314558	AI314558	uj48d07.y
c	37	247.2	47.8	499	12	BI294072	BI294072	UI-R-DK0-
c	38	244.2	47.2	502	9	AI104669	AI104669	EST213958
c	39	243	47.0	561	12	BI714874	BI714874	ic33b09.x
c	40	240.6	46.5	564	12	BI714981	BI714981	ic34h10.x
	41	239.2	46.3	2170	11	AK038119	AK038119	Mus muscu
	42	237.4	45.9	558	12	BI715475	BI715475	ic33c08.y
c	43	237.2	45.9	480	9	AA621551	AA621551	af47c10.s
	44	236.8	45.8	512	9	AI876203	AI876203	uj59b10.y
	45	234.2	45.3	949	14	CB589117	CB589117	AGENCOURT

## ALIGNMENTS

## RESULT 1

CB959991

LOCUS CB959991 796 bp mRNA linear EST 29-APR-2003

DEFINITION AGENCOURT\_13888044 NIH\_MGC\_147 Homo sapiens cDNA clone  
IMAGE:30341081 5', mRNA sequence.

ACCESSION CB959991

VERSION CB959991.1 GI:30216107

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 796)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Dr. Stefan Hansson

cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help  
and advice from Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:<http://image.llnl.gov>

Plate: NDAM371 row: p column: 18

High quality sequence stop: 707.

## FEATURES

source

Location/Qualifiers

1..796

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30341081"

/tissue\_type="Human Placenta"

/lab\_host="DH10B TonA"

/clone\_lib="NIH\_MGC\_147"

/note="Organ: placenta; Vector: pBluescriptR; Site\_1:

all-XhoI; Site\_2: BamH; Oligo-dT primed using primer

5'-TTTTTTTTTTTTTTVN-3', size-selected for average

insert size 2.3 kb and normalized to ROT 5. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIMH/NHGRI, National Institutes of Health). Note: This is  
a NIH\_MGC library."

BASE COUNT 224 a 197 c 191 g 184 t

## ORIGIN

Query Match 66.6%; Score 344.2; DB 14; Length 796;

Best Local Similarity 87.3%; Pred. No. 8.3e-81;

Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;

Qy 1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60  
 |||  
 Db 180 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 239

Qy 61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120  
 |||  
 Db 240 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 299

Qy 121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180  
 |||  
 Db 300 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 359

Qy 181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240  
 |||  
 Db 360 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 419

Qy 241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300  
 |||  
 Db 420 ATGCCCAAGACCCAG----- 434

Qy 301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360  
 |||  
 Db 435 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 490

Qy 361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419  
 |||  
 Db 491 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 550

Qy 420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTC 477  
 |||  
 Db 551 TGCACGAGTTACCTGTAAACTTTGGAACACCTACCAAAAAATAAGTTTGATAACATTTA 610

Qy 478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517  
 |||  
 Db 611 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 651

## RESULT 2

AI503976/c

LOCUS AI503976 558 bp mRNA linear EST 11-MAR-1999

DEFINITION vm43d08.x1 Stratagene mouse diaphragm (#937303) Mus musculus cDNA clone IMAGE:1001007 3' similar to gb:X04482 Mouse mRNA for preproinsulin-like growth factor IB (MOUSE);, mRNA sequence.

ACCESSION AI503976

VERSION AI503976.1 GI:4401827

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 558)

AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.





```

/clone="UI-CF-EC1-abj-k-24-0-UI"
/tissue_type="Lung"
/dev_stage="Adult and Fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EC1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGTGCTTAC.
TAG_LIB=UI-CF-EC1
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG_SEQ=AAGTGCTTAC"

```

```

BASE COUNT      152 a      164 c      169 g      188 t
ORIGIN

```

```

. Query Match          63.9%; Score 330.6; DB 12; Length 673;
Best Local Similarity  86.9%; Pred. No. 3.3e-77;
Matches 453; Conservative 0; Mismatches 14; Indels 54; Gaps 6;

```

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      492 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 433

Qy      61 AGGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db      432 AGGGGG-TTTTATTTTCAGCAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 374

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      373 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 314

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      313 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 254

Qy      241 ATGCCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      ||||||||||||||||
Db      253 ATGCCCCAAGACCCAG----- 239

Qy      301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      238 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 183

Qy      361 GATGTA-GAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      ||||| ||||||||| ||||||||||||| |||| ||||||||||| ||||||||

```

Db 182 GATGTAGGAAGACCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 123

QY 420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTTC 477  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 122 TGCACGAGTTACCTGTTAAACTTTGGAACACCTACCAAAAAATAAGTTTGATAACATTTA 63

QY 478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 62 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 22

#### RESULT 4

AW146128/c

LOCUS AW146128 623 bp mRNA linear EST 10-OCT-2000

DEFINITION um37e10.x1 Sugano mouse embryo mewa Mus musculus cDNA clone  
 IMAGE:2247498 3' similar to gb:X04482 Mouse mRNA for  
 preproinsulin-like growth factor IB (MOUSE);, mRNA sequence.

ACCESSION AW146128

VERSION AW146128.1 GI:6167864

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 623)

AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,  
 Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person  
 ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter  
 ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,  
 Waterston,R. and Wilson,R.

TITLE The WashU-NCI Mouse EST Project 1999

JOURNAL Unpublished

COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:1006958

Seq primer: custom primer used

High quality sequence stop: 499.

#### FEATURES

source

Location/Qualifiers

1..623

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL"

/db\_xref="taxon:10090"

/clone="IMAGE:2247498"

/dev\_stage="embryo, 14 dpc"

/lab\_host="DH10B"

/clone\_lib="Sugano mouse embryo mewa"

/note="Vector: pME18S-FL3; Site\_1: DraIII (CACTGTGTG);

Site\_2: DraIII (CACCATGTG); 1st strand cDNA was primed

with an oligo(dT) primer [ATGTGGCCTTTTTTTTTTTTTTTTTT];

double-stranded cDNA was ligated to a DraIII adaptor



BASE COUNT	123 a	138 c	170 g	191 t	1 others
ORIGIN					

Query Match 63.8%; Score 329.8; DB 9; Length 623;  
Best Local Similarity 80.6%; Pred. No. 5.3e-77;  
Matches 425; Conservative 0; Mismatches 92; Indels 10; Gaps 3;

Qy            1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGGTGGATGCTCTTTCAGTTCCGTGTGTGGAGAC 60  
||| ||| ||| || | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db          541 GGACCAGAGACCCCTTTTCGGGGCTGAGCTGGTGATGCTCTTCAGTTCCGTGTGTGGACCG 482

Qy            61 AGGGGCTTTTATTTCACAACAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db          481 AGGGGCTTTTACTTCAACAAGCCCACAGGCTATGGCTCCAGCATTCGGAGGGCACCTCAG 422

Qy            121 ACAGGCATCGTGGA TGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db          421 ACAGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGACTGGAAATGTAC 362

Qy            181 TGC GCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240  
|| | || | || | || | || | || | || | || | || | || | || | || | || | || | || | || | || | || |  
Db          361 TGTGCCCCACTGAAGCCTACAAAAGCAGCCCGCTCTATCCGTGCCCAGCGCCACACTGAC 302

Qy            241 ATGCCCAAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCA---G 297  
||| ||| ||| ||| ||| ||| | | ||| ||| ||| ||| ||| ||| ||| || | || |  
Db          301 ATGCCCAAAGACTCAGAAGTCCCCGTCCCTATCGACAAACAAGAAAACGAAGCTGCAAAGG 242

Qy            298 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 357  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db          241 AGAAGGAAAGGAAGTACATTTGAAGAACCCAAGTAGAGGAAGTGCAGGAAACAAGACCTA 182

Qy            358 CAGGATGTA-GAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTG 416  
|| | ||| ||| || | ||| || | || | || | || | || | || | || | || | || | || | || | || | || |  
Db          181 CAGAATGTAGGAGGAGCCTCCACGGAGCAGAAAATGCCACATCACCGCAGGATCCTTTG 122

Qy            417 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCA-----AAAAATAAGTTTGATC 470  
|| | || | ||| ||| || | || | || | || | || | || | || | || | || | || | || | || | || | || |  
Db          121 CTGCTTGAGCAACCTGCAAAACATCGAAACCCCTACCAAATAACAATAAAGTCCAATA 62

Qy            471 ACATTTTCAAAGATGGCATTTC CCCCAATGAAATACACAAGTAAACAT 517  
||| ||| ||| ||| || | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db          61 ACATTACAAAGATGGGCATTTC CCCCAATGAAATATACAAGTAAACAT 15

RESULT 5						
AI248089/c						
LOCUS	AI248089	575 bp	mRNA	linear	EST	01-DEC-1998
DEFINITION	qh69f05.x1 Soares_fetal_liver spleen	1NFLS	S1	Homo sapiens	cdNA	

```

ACCESSION      AI248089
VERSION        AI248089.1  GI:3843486
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 575)
AUTHORS        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                Tumor Gene Index
JOURNAL        Unpublished
COMMENT        Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-r@mail.nih.gov
                This clone is available royalty-free through LLNL ; contact the
                IMAGE Consortium (info@image.llnl.gov) for further information.
                Insert Length: 918 Std Error: 0.00
                Seq primer: -40UP from Gibco
                High quality sequence stop: 380.
FEATURES
  source        Location/Qualifiers
                1..575
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:1849953"
                /sex="male"
                /dev_stage="20 week-post conception fetus"
                /lab_host="DH10B (ampicillin resistant)"
                /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                This is a subtracted version of the original Soares fetal
                liver spleen 1NFLS library. 1st strand cDNA was primed
                with a Pac I - oligo(dT) primer [5'
                AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTTTT 3'],
                double-stranded cDNA was ligated to Eco RI adaptors
                (Pharmacia), digested with Pac I and cloned into the Pac I
                and Eco RI sites of the modified pT7T3 vector. Library
                went through one round of normalization. Library
                constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT     135 a      152 c      131 g      156 t      1 others
ORIGIN

```

Query Match 63.2%; Score 326.6; DB 9; Length 575;  
Best Local Similarity 86.6%; Pred. No. 3.7e-76;  
Matches 438; Conservative 0; Mismatches 15; Indels 53; Gaps 5;

Qy 16 TGC GGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGACAGGGGCTTTTATTTC 75  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 551 TGC GGGGCTGAGCTGGTGNATGCTCTTCAGTTTCGTGTGTGAAGACAGGGGCTTTTATTTC 492

Qy 76 AAC AAGCCCA CAGGGTATGGCTCCAGCAGTCCGAGGGCGCCTCAGACAGGCATCGTGGAT 135  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 491 AAC AAGCCCA CAGGGTATGGCTCCAGCAGTCCGAGGGCGCCTCAGACAGGCATCGTGGAT 432

Qy 136 GAGTGTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTATTGCGCACCCCTCAAG 195  
 |||  
 Db 431 GAGTGTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTATTGCGCACCCCTCAAG 372  
 Qy 196 CCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGACATGCCCAAGACCCAG 255  
 |||  
 Db 371 CCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGACATGCCCAAGACCCAG 312  
 Qy 256 AAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGAAGGAAAGGAAGTACA 315  
 |||  
 Db 311 -----AAGGAAGTACA 301  
 Qy 316 TTTGAAGAACACAAGTAGAGGGAGTGAGGAAACAAGAACTACAGGATGTA-GAAGACCC 374  
 |||  
 Db 300 TTTGAAGAACGCAAGTAGAGGGAGTGAGGAAACAAGAACTACAGGATGTAGGAAGACCC 241  
 Qy 375 TTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTCTGCAC-AGTTACCTG 433  
 |||  
 Db 240 TCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTCTGCACGAGTTACCTG 181  
 Qy 434 -TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAAAGAT-GGCATTTTC 491  
 |||  
 Db 180 TTAAACTTTGGAACACCTACCAAAAAAATAAGTTTGATAACATTTAAAAGATGGGCGTTTC 121  
 Qy 492 CCCCAATGAAATACACAAGTAAACAT 517  
 |||  
 Db 120 CCCCAATGAAATACACAAGTAAACAT 95

# RESULT 6

AI169253/c

LOCUS AI169253 549 bp mRNA linear EST 08-JAN-1999  
 DEFINITION EST215088 Normalized rat kidney, Bento Soares Rattus sp. cDNA clone  
 RKIBP33 3' end, mRNA sequence.

ACCESSION AI169253

VERSION AI169253.1 GI:4134375

KEYWORDS EST.

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 549)

AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
 Kerlavage,A.R. and Adams,M.D.

TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat  
 Gene Index

JOURNAL Unpublished

COMMENT On Oct 6, 1998 this sequence version replaced gi:3705561.

Other\_ESTs: TC50779

Contact: Lee, NH

The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA

Tel: (301)-838-3529

Fax: (301)-838-0208

Email: nhlee@tigr.org

Seq primer: M13-21.

FEATURES  
 source  
 Location/Qualifiers  
 1. 549  
 /organism="Rattus sp."  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10118"  
 /clone="RKIBP33"  
 /clone\_lib="Normalized rat kidney, Bento Soares"  
 /note="Organ: kidney; Vector: pT7T3Pac; Site\_1: EcoRI;  
 Site\_2: NotI"

BASE COUNT 112 a 140 c 133 g 164 t  
 ORIGIN

Query Match 61.2%; Score 316.6; DB 9; Length 549;  
 Best Local Similarity 80.8%; Pred. No. 1.7e-73;  
 Matches 421; Conservative 0; Mismatches 89; Indels 11; Gaps 4;

```

QY      8 AGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGACAGGGGCT 67
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      549 AGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAATTTCGTGTGTGGACCAAGGGGCT 490

QY      68 TTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAGACAGGCA 127
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      489 TTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTCGAAGGGCACCACAGACGGGCA 430

QY      128 TCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTATTGCGCAC 187
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      429 TTGTGGATGAGTGCTAGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTACTGTGCTC 370

QY      188 CCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGACATGCCCA 247
      | ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      369 CGCTGAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACATTGACATGCCCA 310

QY      248 AGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCA---GAGAAGGA 304
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      309 AGACTCAGAAGTCCCAGCCCCATCGACACACAAGAAAAGGAAGCTGCAAAGGAGAAGGA 250

QY      305 AAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAGGATG 364
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      249 AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACTTACAGAATG 190

QY      365 TA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTCTGCA 423
      || || || || || || || || || || || || || || || || || || || ||
Db      189 TAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTGCTGCTTG 130

QY      424 CAGTTACCTGTAAACATTTGGAATACCGGCCA-----AAAAATAAGTTTGATCACATTTT 477
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      129 AGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATACCATTTT 70

QY      478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      69 AGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACAT 29
  
```

RESULT 7  
 AI265629/c  
 LOCUS AI265629 558 bp mRNA linear EST 18-NOV-1998  
 DEFINITION uj04b07.x1 Sugano mouse liver mlia Mus musculus cDNA clone

IMAGE:1890901 3' similar to gb:X04482 Mouse mRNA for preproinsulin-like growth factor IB (MOUSE);, mRNA sequence.

ACCESSION AI265629

VERSION AI265629.1 GI:3883787

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 558)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL Unpublished

COMMENT Contact: Marra M/Mouse EST Project  
WashU-HHMI Mouse EST Project  
Washington University School of MedicineP  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:975225  
Seq primer: custom primer used  
High quality sequence stop: 495.

FEATURES

source

Location/Qualifiers

1. .558

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL"

/db\_xref="taxon:10090"

/clone="IMAGE:1890901"

/sex="female"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="Sugano mouse liver mlia"

/note="Organ: liver; Vector: pME18S-FL3; Site\_1: DraIII (CACTGTGTG); Site\_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGACA."

BASE COUNT 106 a 135 c 156 g 161 t

ORIGIN

Query Match 61.1%; Score 315.8; DB 9; Length 558;

Best Local Similarity 80.8%; Pred. No. 2.7e-73;  
Matches 408; Conservative 0; Mismatches 87; Indels 10; Gaps 3;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      506 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGACCG 447

Qy      61 AGGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      446 AGGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTTCGGAGGGCACCTCAG 387

Qy     121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     386 ACAGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGACTGGAGATGTAC 327

Qy     181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240
      || || || || || || || || || || || || || || || || || || || || ||
Db     326 TGTGCCCCACTGAAGCCTACAAAAGCAGCCCGCTCTATCCGTGCCAGCGCCCACTGAC 267

Qy     241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCA---G 297
      ||||| ||||| ||||| || || || || || || || || || || || || || || ||
Db     266 ATGCCCAAGACTCAGAAGTCCCGTCCCTATCGACAAACAAGAAAACGAAGCTGCAAAGG 207

Qy     298 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 357
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     206 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 147

Qy     358 CAGGATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTG 416
      ||| ||||| || || || || || || || || || || || || || || || || || ||
Db     146 CAGAATGTAGGAGGAGCCTCCACGGAGCAGAAAAATGCCACATCACCGCAGGATCCTTTG 87

Qy     417 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCA-----AAAAATAAGTTTGATC 470
      || || || || || || || || || || || || || || || || || || || || ||
Db      86 CTGCTTGAGCAACCTGCAAAACATCGAAACACCTACCAAATAACAATAATAAGTCCAATA 27

Qy     471 ACATTTCAAAGATGGCATTTCCTCCC 495
      ||||| ||||| || || || || || || || || || || || || || || || || ||
Db      26 ACATTACAAAGATGGGCATTTCCTCCC 2
```

#### RESULT 8

AA542914/c

LOCUS AA542914 498 bp mRNA linear EST 19-AUG-1997

DEFINITION ni98c10.s1 NCI\_CGAP\_Pr21 Homo sapiens cDNA clone IMAGE:984882 3'  
similar to gb:X57025\_rnal INSULIN-LIKE GROWTH FACTOR IA PRECURSOR  
(HUMAN);, mRNA sequence.

ACCESSION AA542914

VERSION AA542914.1 GI:2291394

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 498)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index



Db 237 CATGCCCAAGACCCAG----- 222

Qy 300 AAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACA 359  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 221 -----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACA 167

Qy 360 GGATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCT 418  
 ||||||| ||||||| ||||||| ||||||| ||||| ||||||| ||||||| |||||||

Db 166 GGATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCT 107

Qy 419 CTGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT 476  
 ||||||| ||||||| ||||||| ||||||| ||||| ||||||| ||||||| |||||||

Db 106 CTGCACGAGTTACCTGTTAAACTTTGGAACACCTACCAAAAAATAAGTTTGATAACATTT 47

Qy 477 CAAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517  
 ||||||| ||| |||||||||||||||||||||||||||||||||||

Db 46 AAAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 5

# RESULT 9

CD373004

LOCUS CD373004 614 bp mRNA linear EST 29-MAY-2003

DEFINITION UI-R-GR0-csv-j-17-0-UI.r1 UI-R-GR0 Rattus norvegicus cDNA clone  
 UI-R-GR0-csv-j-17-0-UI 5', mRNA sequence.

ACCESSION CD373004

VERSION CD373004.1 GI:31157094

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 614)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: James Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

<http://genome.uiowa.edu/distribution/rat.html>

Seq primer: M13 REVERSE.

## FEATURES

source

Location/Qualifiers

1..614

/organism="Rattus norvegicus"

/mol\_type="mRNA"





Qy 417 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT 476  
 || | ||||| ||||| ||| ||||| | || ||| |||  
 Db 536 CTGCTTGAGCAACCTGCANAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 595

Qy 477 CAA 479  
 ||  
 Db 596 CCA 598

RESULT 10

AI119218

LOCUS AI119218 816 bp mRNA linear EST 02-SEP-1998

DEFINITION ue94h02.y1 Sugano mouse embryo mewa Mus musculus cDNA clone  
 IMAGE:1498803 5' similar to gb:X04482 Mouse mRNA for  
 preproinsulin-like growth factor IB (MOUSE);, mRNA sequence.

ACCESSION AI119218

VERSION AI119218.1 GI:3519542

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 816)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
 Waterston,R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL Unpublished

COMMENT Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:936407

Seq primer: custom primer used

High quality sequence stop: 473.

FEATURES

source Location/Qualifiers  
 1. .816  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:1498803"  
 /dev\_stage="embryo, 14 dpc"  
 /lab\_host="DH10B"  
 /clone\_lib="Sugano mouse embryo mewa"  
 /note="Vector: pME18S-FL3; Site\_1: DraIII (CACTGTGTG);  
 Site\_2: DraIII (CACCATGTG); 1st strand cDNA was primed  
 with an oligo(dT) primer [ATGTGGCCTTTTTTTTTTTTTTTTTT];  
 double-stranded cDNA was ligated to a DraIII adaptor  
 [TGTTGGCCTACTGG], digested and cloned into distinct DraIII

BASE COUNT	230 a	219 c	172 g	187 t	8 others
ORIGIN					

**Qy**            1 GGACCGGAGACGCCTGTGCGGGGCTGAGCTGGTGGATGCTCCTTCAGTTCCGTGTGTGGAGAC 60  
| | | | |  
**Db**          323 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGATGCTCCTTCAGTTCCGTGTGTGGACCG 382

**Qy**            61 AGGGGCTTTTTATTTCACAACGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120  
| | | | |  
**Db**          383 AGGGGCTTTTTACTTCAACAAGCCCACAGGCTATGGCTCCAGCATTCGGAGGGCACCTCAG 442

**Qy**            121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180  
| | | | |  
**Db**          443 ACAGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGACTGGAGATGTAC 502

**Qy**            181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240  
| | | | |  
**Db**          503 TGTGCCCCACTGAAGCCTACAAAAGCAGCCCGCTCTATCCGTGCCAGCGCCACACTGAC 562

**Qy**            241 ATGCCCAAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCA---G 297  
| | | | |  
**Db**          563 ATGCCCAAAGACTCAGAAGTCCCCGTCCCTATCGACAAACAAGAAAAAGAACTGCAAAGG 622

**Qy**            298 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 357  
| | | | |  
**Db**          623 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGAAGTGCANGAAACAAGACCTA 682

**Qy**            358 CAGGATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 416  
| | | | |  
**Db**          683 CAGAATGTANGAGGAGCCTNCCACGGAGCAGAANATGCCACATCACCGCANGATCCTTTG 742

**Qy**            417 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAATAAGTTTGATCACATT 475  
| | | | |  
**Db**          743 CTGCTTGAGCAACCTGCANAACATCGAAACACCTACCAAATAACATNTATAAGTCCAAT 801

```

RESULT 11
BF383724
LOCUS          BF383724                594 bp    mRNA    linear    EST 27-NOV-2000
DEFINITION     602044632F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:4194295 5',
               mRNA sequence.
ACCESSION      BF383724
VERSION        BF383724.1  GI:11365029
KEYWORDS       EST.
SOURCE         Mus musculus (house mouse)

```





tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT        125 a        134 c        119 g        149 t  
ORIGIN

Query Match                    58.0%;    Score 299.8;    DB 9;    Length 527;  
Best Local Similarity    85.5%;    Pred. No. 4.9e-69;  
Matches 413;    Conservative    0;    Mismatches    17;    Indels    53;    Gaps    5;

```

Qy      39 TCTTCAGTTCGTGTGTGGAGACAGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTC 98
      |||
Db      527 TCTTCAGTTCGTGTGTGGAGACAGGGGCTTTATTTACAACAAGCCCACAGGGTATGGCTC 468

Qy      99 CAGCAGTCGGAGGGCGCCTCAGACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGA 158
      |||
Db      467 CAGCAGTCGGAGGGCGCCTAAGACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGA 408

Qy     159 TCTAAGGAGGCTGGAGATGTATTGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGT 218
      |||
Db      407 TCTAAGGAGGCTGGAGATGTATTGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGT 348

Qy     219 CCGTGCCCGAGCGCCACACCGACATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAA 278
      |||
Db      347 CCGTGCCCGAGCGCCACACCGACATGCCCAAGACCCAG----- 311

Qy     279 CAAGAACACGAAGTCTCAGAGAAGGAAAGGAAGTACATTTGAAGAACAAGTAGAGGGA 338
      |||
Db      310 -----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGA 277

Qy     339 GTGCAGGAAACAAGAACTACAGGATGTA-GAAGACCCCTTCTGAGGAGTGAAGAAGGACAG 397
      |||
Db      276 GTGCAGGAAACAAGAACTACAGGATGTAGGAAGACCCCTCCTGAGGAGTGAAGAGTGACAT 217

Qy     398 GCCACCGCAGGACCCCTTTGCTCTGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAA 455
      |||
Db      216 GCCACCGCAGGATCCTTTGCTCTGCACGAGTTACCTGTTAACTTTGGAACACCTACCAA 157

Qy     456 AAAATAAGTTTGATCACATTTCAAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAA 514
      |||
Db      156 AAAATAAGTTTGATAACATTTAAAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAA 97

Qy     515 CAT 517
      |||
Db      96 CAT 94

```

# RESULT 13

AI876493/c

LOCUS            AI876493                    642 bp        mRNA        linear        EST 21-JUL-1999

DEFINITION    uj59b10.x1 Sugano mouse liver mlia Mus musculus cDNA clone  
IMAGE:1924219 3' similar to gb:X57025\_rnal INSULIN-LIKE GROWTH  
FACTOR IA PRECURSOR (HUMAN); gb:X04482 Mouse mRNA for  
preproinsulin-like growth factor IB (MOUSE);, mRNA sequence.

ACCESSION AI876493  
 VERSION AI876493.1 GI:5550542  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 642)  
 AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,  
 Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person  
 ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter  
 ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,  
 Waterston,R. and Wilson,R.  
 TITLE The WashU-NCI Mouse EST Project 1999  
 JOURNAL Unpublished  
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:980511  
 Seq primer: custom primer used  
 High quality sequence stop: 257.  
 FEATURES  
 source Location/Qualifiers  
 1. .642  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:1924219"  
 /sex="female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="Sugano mouse liver mlia"  
 /note="Organ: liver; Vector: pME18S-FL3; Site\_1: DraIII  
 (CACTGTGTG); Site\_2: DraIII (CACCATGTG); 1st strand cDNA  
 was primed with an oligo(dT) primer  
 [ATGTGGCCTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was  
 ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested  
 and cloned into distinct DraIII sites of the pME18S-FL3  
 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should  
 be used to isolate the cDNA insert. Size selection was  
 performed to exclude fragments <1.5kb. Library  
 constructed by Dr. Sumio Sugano (University of Tokyo  
 Institute of Medical Science). Custom primers for  
 sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end  
 primer CGACCTGCAGCTCGAGCACA."  
 BASE COUNT 127 a 154 c 175 g 185 t 1 others  
 ORIGIN  
 Query Match 56.0%; Score 289.6; DB 9; Length 642;  
 Best Local Similarity 78.9%; Pred. No. 2.7e-66;  
 Matches 397; Conservative 0; Mismatches 95; Indels 11; Gaps 4;

Qy 2 GACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGACA 61  
 |||| |||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 503 GACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGATGCTCTTCAGGTTCGTGTGTGGACCGA 444

Qy 62 GGGGCTTTTATTTCAACAAGCCCAAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAGA 121  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 443 GGGGCTTTTCTTCAACAAGGCCACAGGCTATGGCTCCAGCATTTGGAGGGCACCTCAGA 384

Qy 122 CAGGCATCGTGGATGAGTGCTGCTTCCGG-AGCTGTGATCTAAGGAGGCTGGAGATGTAT 180  
 ||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 383 CAGTCAATGTGGATGAGTGTGCTTCCGGAAGCTGTGATCTGAGAAGACTGNAGATGTAC 324

Qy 181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240  
 || || || || |||| || || |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 323 TGTGCCCCACTGAAGCCTACAAAAGCAGCCGCTCTATCCGTGCCAGCGCCACACTGAC 264

Qy 241 ATGCCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCA---G 297  
 |||| |||| |||| || || |||| || || |||| || || |||| || || |||| || ||  
 Db 263 ATGCCCCAAGACTCAGAAGTCCCGTCCCTATCGACAAACAAGAAAACGAAGCTGCAAAGG 204

Qy 298 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 357  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 203 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 144

Qy 358 CAGGATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 416  
 || |||| || || |||| || || |||| || || |||| || || |||| || || ||||  
 Db 143 CAGAATGTAGGAGGAGCCTCCACGGAGCAGAAAATGCCACATCACCGCAGGATCCTTTG 84

Qy 417 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCA-----AAAAATAAGTTTGATC 470  
 || || |||| || || |||| || || |||| || || |||| || || |||| || || ||||  
 Db 83 CTGCTTGAGCAACCTGCAAAACATCGAAACACCTACCAAATAACAATAATAAGTCCAATA 24

Qy 471 ACATTTCAAAGATGGCATTTCCT 493  
 |||| |||| |||| || || ||||  
 Db 23 ACATTACAAAGATGGGCATTTCCT 1

# RESULT 14

AW495481/c

LOCUS AW495481 499 bp mRNA linear EST 24-FEB-2000

DEFINITION UI-M-BH3-ay-g-11-0-UI.s1 NIH\_BMAP\_M\_S4 Mus musculus cDNA clone  
 UI-M-BH3-ay-g-11-0-UI 3', mRNA sequence.

ACCESSION AW495481

VERSION AW495481.1 GI:7065762

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 499)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: Chin, H



National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mEST@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized pineal glands library cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements  
Seq primer: M13 Forward  
POLYA=Yes.

#### FEATURES

source

#### Location/Qualifiers

1. .499  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UI-M-BH3-auy-g-11-0-UI"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NIH\_BMAP\_M\_S4"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; The NIH\_BMAP\_M\_S4 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH\_BMAP\_M\_S4, NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, NIH\_BMAP\_M\_S3.1, NIH\_BMAP\_M\_S2, NIH\_BMAP\_M\_S1. The subtracted library (NIH\_BMAP\_M\_S4) was constructed as follows: PCR amplified cDNA inserts from NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH\_BMAP\_M\_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
TAG\_LIB=NIH\_BMAP\_M\_S4  
TAG\_TISSUE=pineal-glands

TAG\_SEQ=CAGAC"  
BASE COUNT      86 a     112 c     124 g     177 t  
ORIGIN

Query Match                  55.6%;  Score 287.4;  DB 9;  Length 499;  
Best Local Similarity      80.8%;  Pred. No. 9.7e-66;  
Matches 387;  Conservative      0;  Mismatches 81;  Indels 11;  Gaps 4;

```
Qy      50 TGTGTGGAGACAGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGA 109
      |||||
Db      499 TGTGTGGACCGAGGGGCTTTTACTTCAACAAGCCCACAGGCTATGGCTCCAGCATTTCGGA 440

Qy     110 GGGCGCCTCAGACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGC 169
      |||||
Db     439 GGGCACCTCAGACAGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGAC 380

Qy     170 TGGAGATGTATTGCGCACCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGC 229
      |||||
Db     379 TGGAGATGTACTGTGCCCCACTGAAGCCTACAAAAGCAGCCCGCTCTATCCGTGCCCAGC 320

Qy     230 GCCACACCGACATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGA 289
      |||||
Db     319 GCCCACTGACATGCCCAAGACTCAGAAGTCCCCGTCCCTATCGACAAAACAAGAAAACGA 260

Qy     290 AGTCTCA---GAGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGA 346
      ||
Db     259 AGCTGCAAAGGAGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGAAGTGCAGGA 200

Qy     347 AACAAGAACTACAGGATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGC 405
      |||||
Db     199 AACAAGACCTACAGAATGTAGGAGGAGCCTCCACGAGCAGAAAATGCCACATCACCGC 140

Qy     406 AGGACCCTTTGCTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCA-----AAAAA 459
      |||||
Db     139 AGGATCCTTTGCTGCTTGAGCAACCTGCAAAACATCGAAACACCTACCAAATAACAATAA 80

Qy     460 TAAGTTTGATCACATTTCAAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||||
Db      79 TAAGTCCAATAACATTACAAAGATGGGCATTTCCCCCAATGAAATATACAAGTAAACAT 21
```

RESULT 15

AI169770/c

LOCUS          AI169770                  468 bp      mRNA      linear      EST 20-JAN-1999  
DEFINITION     EST215669 Normalized rat liver, Bento Soares Rattus sp. cDNA clone  
                 RLIAT07 3' end, mRNA sequence.

ACCESSION      AI169770

VERSION         AI169770.1  GI:3709810

KEYWORDS        EST.

SOURCE          Rattus sp.

ORGANISM        Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE      1 (bases 1 to 468)

AUTHORS         Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
                 Kerlavage,A.R. and Adams,M.D.

TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat Gene Index  
 JOURNAL Unpublished  
 COMMENT Other\_ESTs: TC50779  
 Contact: Lee, NH  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@tigr.org  
 Seq primer: M13-21.

FEATURES Location/Qualifiers  
 source 1. .468  
 /organism="Rattus sp."  
 /mol\_type="mRNA"  
 /db\_xref="ATCC (inhost):2027570"  
 /db\_xref="taxon:10118"  
 /clone="RLIAT07"  
 /clone\_lib="Normalized rat liver, Bento Soares"  
 /note="Organ: liver; Vector: pT7T3Pac; Site\_1: EcoRI;  
 Site\_2: NotI"

BASE COUNT 85 a 115 c 119 g 149 t  
 ORIGIN

Query Match 53.4%; Score 276; DB 9; Length 468;  
 Best Local Similarity 80.5%; Pred. No. 1e-62;  
 Matches 375; Conservative 0; Mismatches 80; Indels 11; Gaps 4;

```

Qy      63 GGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAGAC 122
      |||
Db      468 GGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTTCGGAGGGCACCACAGAC 409
      |||

Qy      123 AGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTATTG 182
      |||
Db      408 GGGCATTGTGGATGAGTGTGCTCCCGAGCTGTGATCTGAGGAGGTTGGAGATGTACTG 349
      |||

Qy      183 CGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGACAT 242
      |||
Db      348 TGCTCCGCTGAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGACAT 289
      |||

Qy      243 GCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCA---GAG 299
      |||
Db      288 GCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGGAG 229
      |||

Qy      300 AAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACA 359
      |||
Db      228 AAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTACA 169
      |||

Qy      360 GGATGTA-GAAGACCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCTTTGCT 418
      |||
Db      168 GAATGTAGGAGAGCCTCCCAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTGCT 109
      |||

Qy      419 CTGCACAGTTACCTGTAAACATTGGAATACCGGCCA-----AAAAATAAGTTTGATCAC 472
      |||
Db      108 GCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATACC 49
      |||

Qy      473 ATTTCAAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517

```

|||||    |||||    |||||    |||||    |||||

Search completed: December 13, 2003, 07:29:47  
Job time : 1690.26 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 05:41:20 ; Search time 2309.97 Seconds  
(without alignments)  
9156.102 Million cell updates/sec

Title: US-09-852-261-1  
Perfect score: 517  
Sequence: 1 ggaccggagacgctctgcgg.....tgaaatacacacaagtaaacaat 517

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*

28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query Match	Length			
1	517	100.0	517	6	AX147742	AX147742 Sequence
2	517	100.0	517	6	AX300779	AX300779 Sequence
3	467.4	90.4	523	6	AX147746	AX147746 Sequence
4	467.4	90.4	523	6	AX300783	AX300783 Sequence
5	377.2	73.0	471	6	AX147754	AX147754 Sequence
6	377.2	73.0	471	6	AX300791	AX300791 Sequence
7	355.4	68.7	444	9	HSU40870	U40870 Human alter
8	344.2	66.6	616	9	HSIGF1A	X56773 H.sapiens m
9	344.2	66.6	7260	6	AX375028	AX375028 Sequence
10	344.2	66.6	7260	6	AX411095	AX411095 Sequence
11	344.2	66.6	7260	9	HSIGFACI	X57025 Human IGF-I
12	342.6	66.3	666	6	A29119	A29119 H.sapiens I
13	342.6	66.3	725	9	HSIGFI	X00173 Homo sapien
14	342.6	66.3	728	9	HUMGFII	M29644 Human insul
15	342.6	66.3	1076	9	HUMIGFI	M27544 Human insul
16	341	66.0	620	6	I08370	I08370 Sequence 2
17	331.6	64.1	1536	10	BC012409	BC012409 Mus muscu
18	330	63.8	798	10	RNIGFI2	X06108 Rat mRNA (c
19	330	63.8	958	10	RNIGFI1	X06107 Rat mRNA (c
20	326.8	63.2	710	10	RATIGFIA	M15480 Rat insulin
21	325.2	62.9	539	6	AX147744	AX147744 Sequence
22	325.2	62.9	539	6	AX300781	AX300781 Sequence
23	318.2	61.5	651	10	MMIGFIBR	X04482 Mouse mRNA
24	308.6	59.7	730	9	HSIGF1B	X56774 H.sapiens m
25	308.6	59.7	1094	9	HUMGFIB	M11568 Human insul
26	308.6	59.7	1136	6	E01349	E01349 cDNA encodi
27	308.6	59.7	1136	6	I08009	I08009 Sequence 3
28	294.4	56.9	432	4	AF022961	AF022961 Oryctolag
29	286.4	55.4	3599	6	BD063790	BD063790 Insulin-l
30	286.4	55.4	3599	6	BD069041	BD069041 Treatment
31	286.4	55.4	3600	6	BD063789	BD063789 Insulin-l
32	286.4	55.4	3600	6	BD069040	BD069040 Treatment
33	285.4	55.2	612	9	HUMIGFIB	M37484 Human insul

34	283.4	54.8	888	4	ECU85272	U85272 Equus cabal
35	281.2	54.4	532	4	SSILGF1M	X17492 Porcine mRN
36	281.2	54.4	567	4	PIGGFIIA	M31175 Pig insulin
37	276	53.4	1284	4	BTILGF1A	X15726 Bovine mRNA
38	271.2	52.5	978	4	GOTIGFI	D11378 Goat mRNA f
39	271.2	52.5	978	6	E05279	E05279 DNA encodin
40	267.6	51.8	836	10	CPIGF1	X52951 Guinea pig
41	264.8	51.2	747	4	SHPIGFIA6	M31735 Sheep insul
42	264.8	51.2	790	4	SHPIGFIA21	M31734 Sheep insul
43	264.8	51.2	1015	4	SHPIGFIA46	M31736 Sheep insul
44	264.8	51.2	1107	4	SHPIGFIA1	M30653 Sheep insul
45	258.4	50.0	317	6	AX147750	AX147750 Sequence

# ALIGNMENTS

## RESULT 1

AX147742

LOCUS AX147742 517 bp DNA linear PAT 31-AUG-2001

DEFINITION Sequence 1 from Patent W00136483.

ACCESSION AX147742

VERSION AX147742.1 GI:14346787

## KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE 1

AUTHORS Goldspink, G.R. and Johnson, I.R.

TITLE Use of the insulin-like-growth factor i isoform mgf for the treatment of neurological disorders

JOURNAL Patent: WO 0136483-A 1 25-MAY-2001; University College London (GB)

## FEATURES

source Location/Qualifiers

1..517

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

CDS

<1..333

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC41175.1"

/db\_xref="GI:14346788"

/db\_xref="REMTREMBL:CAC41175"

/translation="GPETLCGAELVDALQFVCGDRGFYFNKPTGYGSSRRAPQTGIV  
DECCFRSCDLRRLEMYCAPLKPAKSARSVRAQRHTDMPKTQKYQPPSTNKNTKSQRRK  
GSTFEEHK"

BASE COUNT 150 a 130 c 139 g 98 t

## ORIGIN

Query Match 100.0%; Score 517; DB 6; Length 517;

Best Local Similarity 100.0%; Pred. No. 4.4e-155;

Matches 517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

|||||

Db 1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

Qy 61 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120  
 |||

Db 61 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120

Qy 121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180  
 |||

Db 121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180

Qy 181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240  
 |||

Db 181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240

Qy 241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300  
 |||

Db 241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300

Qy 301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAAGTACAG 360  
 |||

Db 301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAAGTACAG 360

Qy 361 GATGTAGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTCT 420  
 |||

Db 361 GATGTAGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTCT 420

Qy 421 GCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA 480  
 |||

Db 421 GCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA 480

Qy 481 GATGGCATTTCCCCCAATGAAATACACAAGTAAACAT 517  
 |||

Db 481 GATGGCATTTCCCCCAATGAAATACACAAGTAAACAT 517

# RESULT 2

AX300779

LOCUS AX300779 517 bp DNA linear PAT 30-NOV-2001

DEFINITION Sequence 1 from Patent WO0185781.

ACCESSION AX300779

VERSION AX300779.1 GI:17382060

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Goldspink, G.D. and Terenghi, G.B.

TITLE Repair of nerve damage

JOURNAL Patent: WO 0185781-A 1 15-NOV-2001; University College London (GB) ; East Grinstead Medical Research Trust (GB)

FEATURES

source

Location/Qualifiers

1. .517

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"





VERSION AX147746.1 GI:14346791  
 KEYWORDS  
 SOURCE Oryctolagus cuniculus (rabbit)  
 ORGANISM Oryctolagus cuniculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 REFERENCE 1  
 AUTHORS Goldspink, G.R. and Johnson, I.R.  
 TITLE Use of the insulin-like-growth factor i isoform mgf for the  
 treatment of neurological disorders  
 JOURNAL Patent: WO 0136483-A 5 25-MAY-2001;  
 University College London (GB)  
 FEATURES Location/Qualifiers  
 source 1..523  
 /organism="Oryctolagus cuniculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9986"  
 CDS <1..336  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC41177.1"  
 /db\_xref="GI:14346792"  
 /db\_xref="REMTREMBL:CAC41177"  
 /translation="GPETLCGAELVDALQFVCGDRGFYFNKPTGYGSSSSRRAPQTGIV  
 DECCFRSCDLRRLEMYCAPLKPAAARSVRAQRHTDMPKTQKYQPPSTNKKMKSQRRR  
 KGSTFEEHK"  
 BASE COUNT 154 a 129 c 142 g 98 t  
 ORIGIN

Query Match 90.4%; Score 467.4; DB 6; Length 523;  
 Best Local Similarity 96.2%; Pred. No. 4.4e-139;  
 Matches 501; Conservative 0; Mismatches 16; Indels 4; Gaps 2;

Qy	1	GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC	60
Db	1	GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC	60
Qy	61	AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG	120
Db	61	AGGGGCTTTTATTTCAACAAGCCACAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG	120
Qy	121	ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT	180
Db	121	ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC	180
Qy	181	TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC	240
Db	181	TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCCAGCGCCACACCGAC	240
Qy	241	ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCA---G	297
Db	241	ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAATGAAGTCTCAGAGG	300
Qy	298	AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA	357
Db	301	AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA	360

Qy 358 CAGGATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 416  
 |||  
 Db 361 CAGGATGTAGGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 420  
 Qy 417 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT 476  
 |||  
 Db 421 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT 480  
 Qy 477 CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 517  
 |||  
 Db 481 CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 521

#### RESULT 4

AX300783

LOCUS AX300783 523 bp DNA linear PAT 30-NOV-2001

DEFINITION Sequence 5 from Patent WO0185781.

ACCESSION AX300783

VERSION AX300783.1 GI:17382064

KEYWORDS

SOURCE Oryctolagus cuniculus (rabbit)

ORGANISM Oryctolagus cuniculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

REFERENCE 1

AUTHORS Goldspink, G.D. and Terenghi, G.B.

TITLE Repair of nerve damage

JOURNAL Patent: WO 0185781-A 5 15-NOV-2001;

University College London (GB) ; East Grinstead Medical Research  
 Trust (GB)

FEATURES

source

Location/Qualifiers

1. 523

/organism="Oryctolagus cuniculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:9986"

CDS

<1. 336

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD13042.1"

/db\_xref="GI:17382065"

/translation="GPETLCGAELVDALQFVCGDRGFYFNKPTGYGSSRRAPQTGIV  
 DECCFRSCDLRRLEMYCAPLKPAAARSVRAQRHTDMPKTQKYQPPSTNKKMKSQRRR  
 KGSTFEEHK"

BASE COUNT 154 a 129 c 142 g 98 t

ORIGIN

Query Match 90.4%; Score 467.4; DB 6; Length 523;

Best Local Similarity 96.2%; Pred. No. 4.4e-139;

Matches 501; Conservative 0; Mismatches 16; Indels 4; Gaps 2;

Qy 1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60  
 |||  
 Db 1 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60  
 Qy 61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120  
 |||  
 Db 61 AGGGGCTTTTATTTCAACAAGCCACAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG 120

Qy 121 ACAGGCATCGTGGATGAGTGCCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180  
 |||  
 Db 121 ACAGGCATCGTGGATGAGTGCCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy 181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240  
 ||  
 Db 181 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGTCCGTCCGTGCCCAGCGCCACACCGAC 240

Qy 241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCA---G 297  
 |||  
 Db 241 ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAATGAAGTCTCAGAGG 300

Qy 298 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAAACAAGAACTA 357  
 |||  
 Db 301 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAAACAAGAACTA 360

Qy 358 CAGGATGTA-GAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTG 416  
 |||  
 Db 361 CAGGATGTAGGAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTG 420

Qy 417 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT 476  
 |||  
 Db 421 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT 480

Qy 477 CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 517  
 |||  
 Db 481 CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 521

# RESULT 5

AX147754

LOCUS AX147754 471 bp DNA linear PAT 08-JUN-2001

DEFINITION Sequence 13 from Patent WO0136483.

ACCESSION AX147754

VERSION AX147754.1 GI:14348552

KEYWORDS

SOURCE Oryctolagus cuniculus (rabbit)

ORGANISM Oryctolagus cuniculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

REFERENCE 1

AUTHORS Goldspink, G.R. and Johnson, I.R.

TITLE Use of the insulin-like-growth factor i isoform mgf for the  
 treatment of neurological disorders

JOURNAL Patent: WO 0136483-A 13 25-MAY-2001;  
 University College London (GB)

FEATURES

source

Location/Qualifiers

1..471

/organism="Oryctolagus cuniculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:9986"

CDS

<1..318

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC41264.1"

/db\_xref="GI:14348553"

/translation="GPETLCGAELVDALQFVCGDRGFYFNKPTGYGSSSRRAPQTGIV  
DECCFRSCDLRRLEMYCAPLKPAAARSVRAQRHTDMPKTQKEVHLKNTSRGSAGNKN  
YRM"

BASE COUNT 132 a 118 c 131 g 90 t  
ORIGIN

Query Match 73.0%; Score 377.2; DB 6; Length 471;  
Best Local Similarity 87.8%; Pred. No. 5.4e-110;  
Matches 455; Conservative 0; Mismatches 13; Indels 50; Gaps 2;

```
QY      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db      1 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

QY     61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db     61 AGGGGCTTTTATTTCAACAAGCCACAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG 120

QY    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

QY    181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      ||
Db    181 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCCAGCGCCACACCGAC 240

QY    241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db    241 ATGCCCAAGACTCAG----- 255

QY    301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db    256 ----AAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 311

QY    361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db    312 GATGTAGGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 371

QY    420 TGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA 479
      |||
Db    372 TGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA 431

QY    480 AGATGGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||
Db    432 AGATGGCATTTCCCCCAATGAAATACACAAGTAAACAT 469
```

RESULT 6  
AX300791

LOCUS AX300791 471 bp DNA linear PAT 30-NOV-2001  
DEFINITION Sequence 13 from Patent WO0185781.  
ACCESSION AX300791  
VERSION AX300791.1 GI:17382072  
KEYWORDS  
SOURCE Oryctolagus cuniculus (rabbit)  
ORGANISM Oryctolagus cuniculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

REFERENCE 1

AUTHORS Goldspink, G.D. and Terenghi, G.B.

TITLE Repair of nerve damage

JOURNAL Patent: WO 0185781-A 13 15-NOV-2001;  
University College London (GB) ; East Grinstead Medical Research  
Trust (GB)

FEATURES Location/Qualifiers

source 1. .471  
/organism="Oryctolagus cuniculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9986"

CDS <1. .318  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD13045.1"  
/db\_xref="GI:17382073"  
/translation="GPETLCGAELVDALQFVCGDRGFYFNKPTGYGSSSSRRAPQTGIV  
DECCFRSCDLRRLLEMYCAPLKPAAKARSVRAQRHTDMPKTQKEVHLKNTSRGSAGNKN  
YRM"

BASE COUNT 132 a 118 c 131 g 90 t

ORIGIN

Query Match 73.0%; Score 377.2; DB 6; Length 471;  
Best Local Similarity 87.8%; Pred. No. 5.4e-110;  
Matches 455; Conservative 0; Mismatches 13; Indels 50; Gaps 2;

Qy 1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 60  
|||||

Db 1 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 60

Qy 61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120  
|||||

Db 61 AGGGGCTTTTATTTCAACAAGCCACAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG 120

Qy 121 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180  
|||||

Db 121 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy 181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240  
|||

Db 181 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCCAGCGCCACACCGAC 240

Qy 241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300  
|||||

Db 241 ATGCCCAAGACTCAG----- 255

Qy 301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360  
|||||

Db 256 ----AAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 311

Qy 361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419  
|||||

Db 312 GATGTAGGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 371

Qy 420 TGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAATAAGTTTGATCACATTTCAA 479  
|||||

Db 372 TGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA 431

Qy 480 AGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 517

Db 432 AGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 469

# RESULT 7

HSU40870

LOCUS HSU40870 444 bp mRNA linear PRI 05-APR-1996

DEFINITION Human alternatively spliced human insulin-like growth factor-I (IGF-I) mRNA, partial cds.

ACCESSION U40870

VERSION U40870.1 GI:1100902

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 444)

AUTHORS Chew, S.L., Lavender, P., Clark, A.J. and Ross, R.J.

TITLE An alternatively spliced human insulin-like growth factor-I transcript with hepatic tissue expression that diverts away from the mitogenic IBE1 peptide

JOURNAL Endocrinology 136 (5), 1939-1944 (1995)

MEDLINE 95237119

PUBMED 7720641

REFERENCE 2 (bases 1 to 444)

AUTHORS Chew, S.L.

TITLE Direct Submission

JOURNAL Submitted (20-NOV-1995) Shern L. Chew, Endocrinology, St Bartholomew's Hospital Medical College, West Smithfield, London, EclA 7Be, UK

FEATURES Location/Qualifiers

source

1..444  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="pC4"  
/tissue\_type="liver"

gene

1..444  
/gene="IGF-I"

CDS

<1..420  
/gene="IGF-I"  
/note="alternatively spliced; previously, exon 5 and 6 were thought to be mutually exclusive; this transcript splices from exon 5 into exon 6; the alternatively spliced transcript would continue with exon 5 to the polyA signal"  
/codon\_start=1  
/product="insulin-like growth factor-I"  
/protein\_id="AAA96152.1"  
/db\_xref="GI:1100903"  
/translation="LKVKMHTMSSSHLFYLA LCLLTFTSSATAGPETLCGAELVDALQ FVCGDRGFYFNKPTGYGSSRRAPQTGIVDECCFRSCDLRRLEMYCAPLKPASARSV RAQRHTDMPKTQKYQPPSTNKNTKSQRKRGSTFEERK"

exon

1..6  
/gene="IGF-I"

```

                /number=1
    exon         7. .163
                /gene="IGF-I"
                /number=3
    exon         164. .345
                /gene="IGF-I"
                /number=4
    exon         346. .394
                /gene="IGF-I"
                /number=5
    exon         395. .420
                /gene="IGF-I"
                /number=6
BASE COUNT      107 a    125 c    125 g    87 t
ORIGIN

```

Query Match 68.7%; Score 355.4; DB 9; Length 444;  
 Best Local Similarity 99.7%; Pred. No. 5.7e-103;  
 Matches 356; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 147

Qy     61 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     148 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 207

Qy     121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 267

Qy     181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 327

Qy     241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 387

Qy     301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 357
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     388 AGGAAAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTA 444

```

# RESULT 8

HSIGF1A

LOCUS HSIGF1A 616 bp mRNA linear PRI 29-NOV-1993

DEFINITION H.sapiens mRNA for IGF-1a.

ACCESSION X56773 S61841

VERSION X56773.1 GI:32989

KEYWORDS IGF-1 gene.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 616)



AUTHORS Sandberg-Nordqvist,A.C., Stahlbom,P.A., Lake,M. and Sara,V.R.  
 TITLE Characterization of two cDNAs encoding insulin-like growth factor 1 (IGF-1) in the human fetal brain  
 JOURNAL Brain Res. Mol. Brain Res. 12 (1-3), 275-277 (1992)  
 MEDLINE 92186627  
 PUBMED 1372070  
 REFERENCE 2 (bases 1 to 616)  
 AUTHORS Sandberg Nordqvist,A.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-NOV-1990) A.C.Sandberg Nordqvist, KAROLINSKA INST'S DEPT OF PATHOLOGY, KAROLINSKA HOSPITAL, BOX 605 00, S-104 01 STOCKHOLM, SWEDEN  
 REFERENCE 3 (bases 1 to 616)  
 AUTHORS Sandberg-Nordqvist,A.C., Stahlbom,P.A., Reinecke,M., Collins,V.P., von Holst,H. and Sara,V.  
 TITLE Characterization of insulin-like growth factor 1 in human primary brain tumors  
 JOURNAL Cancer Res. 53 (11), 2475-2478 (1993)  
 MEDLINE 93265440  
 PUBMED 8495408

FEATURES Location/Qualifiers  
 source 1. .616  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /map="q22-q24"  
 /tissue\_type="brain"  
 /dev\_stage="fetal"  
 gene 1. .462  
 /gene="IGF-1"  
 CDS 1. .462  
 /gene="IGF-1"  
 /codon\_start=1  
 /product="IGF-1a"  
 /protein\_id="CAA40092.1"  
 /db\_xref="GI:32990"  
 /db\_xref="SWISS-PROT:P01343"  
 /translation="MGKISSLP TQLFKCCFCDFLKVKMHTMSSSHLFYLALCLLTFTS  
 SATAGPETLCGAELVDALQFVCGDRGFYFNKPTGYGSSSRAPQTGIVDECCFRSCDL  
 RRLEMYCAPLKPAKSARSVRAQRHTDMPKTQKEVHLKNASRGSAKNYRM"  
 mat\_peptide 145. .354  
 /gene="IGF-1"  
 /product="IGF-1a"  
 exon 403. .616  
 /note="exon 5"  
 BASE COUNT 159 a 158 c 160 g 139 t  
 ORIGIN

Query Match 66.6%; Score 344.2; DB 9; Length 616;  
 Best Local Similarity 87.3%; Pred. No. 2.5e-99;  
 Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;

Qy. 1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60  
 |||||  
 Db 145 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 204

Qy 61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120  
 |||  
 Db 205 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 264  
 Qy 121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180  
 |||  
 Db 265 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 324  
 Qy 181 TGCGCACCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240  
 |||  
 Db 325 TGCGCACCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 384  
 Qy 241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300  
 |||  
 Db 385 ATGCCCAAGACCCAG----- 399  
 Qy 301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360  
 |||  
 Db 400 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 455  
 Qy 361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419  
 |||  
 Db 456 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 515  
 Qy 420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTT 477  
 |||  
 Db 516 TGCACGAGTTACCTGTTAACTTTGGAACACCTACCAAAAAATAAGTTTGATAACATTTA 575  
 Qy 478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517  
 |||  
 Db 576 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 616

RESULT 9

AX375028

LOCUS AX375028 7260 bp DNA linear PAT 01-MAR-2002

DEFINITION Sequence 31 from Patent WO0210436.

ACCESSION AX375028

VERSION AX375028.1 GI:19169860

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Baak,J. and Mutter,G.L.

TITLE Prognostic classification of breast cancer

JOURNAL Patent: WO 0210436-A 31 07-FEB-2002;

THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; Baak, Jan (US)

FEATURES Location/Qualifiers

source 1. 7260

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 2330 a 1415 c 1240 g 2275 t

ORIGIN

Query Match 66.6%; Score 344.2; DB 6; Length 7260;  
 Best Local Similarity 87.3%; Pred. No. 3.6e-99;  
 Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db      311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 370

Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db      371 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 430

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db      431 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240
      |||
Db      491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 550

Qy      241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db      551 ATGCCCAAGACCCAG----- 565

Qy      301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db      566 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 621

Qy      361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db      622 GATGTAGGAAGACCCTCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 681

Qy      420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTTC 477
      |||
Db      682 TGCACGAGTTACCTGTTAAACTTTGGAACACCTACCAAAAAATAAGTTTGATAACATTTA 741

Qy      478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||
Db      742 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 782
  
```

# RESULT 10

AX411095

LOCUS AX411095 7260 bp DNA linear PAT 14-JUN-2002

DEFINITION Sequence 3742 from Patent WO0229103.

ACCESSION AX411095

VERSION AX411095.1 GI:21443800

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.

TITLE Gene expression profiles in liver cancer

JOURNAL Patent: WO 0229103-A 3742 11-APR-2002;

GENE LOGIC INC (US)

FEATURES                      Location/Qualifiers  
     source                    1. .7260  
                               /organism="Homo sapiens"  
                               /mol\_type="genomic DNA"  
                               /db\_xref="taxon:9606"  
                               /note="EMBL/GenBank Accession No. X57025"

BASE COUNT        2330 a    1415 c    1240 g    2275 t  
 ORIGIN

Query Match                      66.6%;    Score 344.2;    DB 6;    Length 7260;  
 Best Local Similarity        87.3%;    Pred. No. 3.6e-99;  
 Matches 455;    Conservative    0;    Mismatches    13;    Indels    53;    Gaps    5;

```

Qy      1  GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db    311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 370

Qy      61  AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db    371  AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 430

Qy     121  ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    431  ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490

Qy     181  TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240
      |||
Db    491  TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 550

Qy     241  ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db    551  ATGCCCAAGACCCAG----- 565

Qy     301  AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db    566  ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 621

Qy     361  GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db    622  GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 681

Qy     420  TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTC 477
      |||
Db    682  TGCACGAGTTACCTGTTAAACTTTGGAACACCTACCAAAAAATAAGTTTGATAACATTTA 741

Qy     478  AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||
Db    742  AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 782
  
```

# RESULT 11

HSIGFACI

LOCUS            HSIGFACI                      7260 bp    mRNA    linear    PRI 17-FEB-1992

DEFINITION    Human IGF-I mRNA for insulin-like growth factor I.

ACCESSION    X57025

VERSION      X57025.1    GI:33007

KEYWORDS     insulin-like growth factor I.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 7260)

AUTHORS Steenbergh, P.H., Koonen-Reemst, A.M., Cleutjens, C.B. and  
Sussenbach, J.S.

TITLE Complete nucleotide sequence of the high molecular weight human  
IGF-I mRNA

JOURNAL Biochem. Biophys. Res. Commun. 175 (2), 507-514 (1991)

MEDLINE 91207342

PUBMED 2018498

REFERENCE 2 (bases 1 to 7260)

AUTHORS Steenbergh, P.H.

TITLE Direct Submission

JOURNAL Submitted (18-DEC-1990) P.H. Steenbergh, LAB FOR PHYSIOLOGICAL  
CHEMISTRY, UNIVERSITY OF UTRECHT, VONDELLAAN 24 A, 3521 GG UTRECHT,  
THE NETHERLANDS

FEATURES Location/Qualifiers

source 1..7260  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="12 q22-24.1"  
/tissue\_type="liver"  
/dev\_stage="adult"

gene 1..7260  
/gene="IGF-I"

mRNA 1..7260  
/gene="IGF-I"  
/evidence=experimental

exon 1..229  
/gene="IGF-I"  
/number=1  
/evidence=experimental

CDS 167..628  
/gene="IGF-I"  
/codon\_start=1  
/evidence=experimental  
/product="insulin-like growth factor I"  
/protein\_id="CAA40342.1"  
/db\_xref="GI:33008"  
/db\_xref="SWISS-PROT:P01343"  
/translation="MGKISSLPTQLFKCCFCDFLKVKMHTMSSSHLFYLALCLLTFTS  
SATAGPETL CGAELVDALQFVCGDRGFYFNKPTGYGSSRRAPQTGIVDECCFRSCDL  
RRLEMYCAPLKPAKSARSVRAQRHTDMPKTQKEVHLKNASRGSGAGNKNYRM"

sig\_peptide 167..310  
/gene="IGF-I"  
/evidence=experimental

mat\_peptide 311..520  
/gene="IGF-I"  
/product="insulin-like growth factor I"  
/evidence=experimental

exon 230..386  
/gene="IGF-I"  
/number=2  
/evidence=experimental

```

exon          387. .568
               /gene="IGF-I"
               /number=3
               /evidence=experimental
exon          569. .7236
               /gene="IGF-I"
               /number=5
               /evidence=experimental
polyA_signal  861. .865
               /gene="IGF-I"
               /note="1.1 kb mRNA"
               /evidence=experimental
repeat_unit   3986. .4026
               /gene="IGF-I"
               /note="CA-repeat"
               /evidence=experimental
repeat_unit   5926. .6215
               /gene="IGF-I"
               /evidence=experimental
               /rpt_family="AluI"
polyA_signal  7205. .7210
               /gene="IGF-I"
               /note="7.6 kb mRNA"
               /evidence=experimental
BASE COUNT    2330 a   1415 c   1240 g   2275 t
ORIGIN

```

```

Query Match          66.6%; Score 344.2; DB 9; Length 7260;
Best Local Similarity 87.3%; Pred. No. 3.6e-99;
Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;

```

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
          |||
Db      311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 370
          |||
Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
          |||
Db      371 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 430
          |||
Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
          |||
Db      431 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490
          |||
Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
          |||
Db      491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 550
          |||
Qy      241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
          |||
Db      551 ATGCCCAAGACCCAG----- 565
          |||
Qy      301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
          |||
Db      566 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 621
          |||
Qy      361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
          |||

```

Db 622 GATGTAGGAAGACCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 681

Qy 420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTC 477  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 682 TGCACGAGTTACCTGTAAACTTTGGAACACCTACCAAAAAATAAGTTTGATAACATTTA 741

Qy 478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 742 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 782

# RESULT 12

A29119

LOCUS A29119 666 bp DNA linear PAT 15-JUN-1995

DEFINITION H.sapiens IGF1 gene fragment from patent GB2241703.

ACCESSION A29119

VERSION A29119.1 GI:1247520

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 666)

AUTHORS

JOURNAL Patent: GB 2241703-A 3 11-SEP-1991;

FEATURES Location/Qualifiers

source 1. .666

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

CDS 25. .384

/partial

/codon\_start=1

/product="IGF-1 precursor"

/protein\_id="CAA01955.1"

/db\_xref="GI:4529932"

/translation="MALCLLTFTSSATAGPETLCGAELVDALQFVCGDRGFYFNKPTG  
 YGSSRRAPQTGI VDECCFRSCDLRRLEMYCAPLKPASRSVRAQRHTDMPKTQKEV  
 HLKNASRGSAGNKNYRM"

mat\_peptide 67. .276

/product="IGF-1"

BASE COUNT 173 a 167 c 181 g 145 t

ORIGIN

Query Match 66.3%; Score 342.6; DB 6; Length 666;

Best Local Similarity 87.1%; Pred. No. 8.2e-99;

Matches 454; Conservative 0; Mismatches 14; Indels 53; Gaps 5;

Qy 1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 67 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 126

Qy 61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 127 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 186

Qy 121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180

Db	187	ACAGGTATCGTGGATGAGTGTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT	246
Qy	181	TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC	240
Db	247	TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC	306
Qy	241	ATGCCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA	300
Db	307	ATGCCCCAAGACCCAG-----	321
Qy	301	AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG	360
Db	322	----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG	377
Qy	361	GATGTA-GAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTGCTC	419
Db	378	GATGTAGGAAGACCCCTCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC	437
Qy	420	TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAATAAGTTTGATCACATTTTC	477
Db	438	TGCACGAGTTACCTGTAAACTTTGGAACACCTACCAAAAATAAGTTTGATAACATTTA	497
Qy	478	AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT	517
Db	498	AAAGATGGGCGT'TTCCCCCAATGAAATACACAAGTAAACAT	538



```

CDS          12. .473
              /gene="IGF-1"
              /codon_start=1
              /product="insulin-like growth factor 1A precursor"
              /protein_id="CAA24998.1"
              /db_xref="GI:33016"
              /db_xref="SWISS-PROT:P01343"
              /translation="MGKISSLPTQLFKCCFCDFLKVKMHTMSSSHLFYLA LCLLTFTS
polyA_site    725
              /gene="IGF-1"
BASE COUNT    190 a    174 c    183 g    178 t
ORIGIN

Query Match          66.3%; Score 342.6; DB 9; Length 725;
Best Local Similarity 87.1%; Pred. No. 8.3e-99;
Matches 454; Conservative 0; Mismatches 14; Indels 53; Gaps 5;

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db      156 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 215

Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db      216 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 275

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db      276 ACAGGTATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 335

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      |||
Db      336 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 395

Qy      241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db      396 ATGCCCAAGACCCAG----- 410

Qy      301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db      411 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 466

Qy      361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db      467 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 526

Qy      420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTTC 477
      |||
Db      527 TGCACGAGTTACCTGTAAACTTTGGAACACCTACCAAAAAATAAGTTTGATAACATTTA 586

Qy      478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||
Db      587 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 627

```

HUMGFII  
LOCUS HUMGFII 728 bp mRNA linear PRI 08-NOV-1994  
DEFINITION Human insulin-like growth factor I mRNA, complete cds.  
ACCESSION M29644  
VERSION M29644.1 GI:183119  
KEYWORDS insulin-like growth factor.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 728)  
AUTHORS Rall,L.B., Scott,J. and Bell,G.I.  
TITLE Human insulin-like growth factor I and II messenger RNA: isolation  
of complementary DNA and analysis of expression  
JOURNAL Meth. Enzymol. 146, 239-248 (1987)  
MEDLINE 88065102  
PUBMED 3683205  
COMMENT Original source text: Human (adult) liver, cDNA to mRNA.  
FEATURES  
Location/Qualifiers  
source 1..728  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="12q23"  
gene 1..728  
/gene="IGF1"  
CDS 81..473  
/gene="IGF1"  
/note="insulin-like growth factor I precursor"  
/codon\_start=1  
/protein\_id="AAA52543.1"  
/db\_xref="GI:183120"  
/db\_xref="GDB:G00-120-081"  
/translation="MHTMSSSHLFYLALCLLTFTSSATAGPETLCGAELVDALQFVCG  
DRGFYFNKPTGYGSSRRAPQTGIVDECCFRSCDLRRLEMYCAPLKPAKSARSVRAQR  
HTDMPKTQKEVHLKNASRGSAGNKNYRM"  
sig\_peptide 81..155  
/gene="IGF1"  
/note="insulin-like growth factor I signal peptide"  
mat\_peptide 156..365  
/gene="IGF1"  
/product="insulin-like growth factor I"  
BASE COUNT 193 a 174 c 183 g 178 t  
ORIGIN Chromosome 12q23.

Query Match 66.3%; Score 342.6; DB 9; Length 728;  
Best Local Similarity 87.1%; Pred. No. 8.3e-99;  
Matches 454; Conservative 0; Mismatches 14; Indels 53; Gaps 5;

QY 1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60  
|||||  
Db 156 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 215  
QY 61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120  
|||||  
Db 216 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 275

[illegible]

RESULT 15

HUMIGFI

LOCUS	HUMIGFI	1076 bp	mRNA	linear	PRI 08-NOV-1994
-------	---------	---------	------	--------	-----------------

DEFINITION Human insulin-like growth factor mRNA, complete cds.

ACCESSION M27544

VERSION M27544.1 GI:184829

KEYWORDS      insulin-like growth factor.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1076)

AUTHORS Le Bouc, Y., Dreyer, D., Jaeger, F., Binoux, M. and Sondermeyer, P.

TITLE Complete characterization of the human IGF-I nucleotide sequence isolated from a newly constructed adult liver cDNA library

JOURNAL FEBS Lett. 196 (1), 108-112 (1986)

MEDLINE 86108910

PUBMED 2935423

COMMENT Original source text: Human liver, cDNA to mRNA, clones  
lanbda-TG[03,04,05].

FEATURES Location/Qualifiers

```
source      1.  .1076
```

```
/organism="Homo sapiens"
```

```
/mol type="mRNA"
```

```
/db xref="taxon:9606"
```

/map="7p13-p12"

gene 1. .1076

```
/gene="IGFBP1"
```

```

mRNA      <1. .>1076
          /gene="IGFBP1"
          /note="IGF mRNA (alt.)"
mRNA      <1. .989
          /gene="IGFBP1"
          /note="IGF mRNA (alt.)"
CDS       149. .610
          /gene="IGFBP1"
          /note="insulin-like growth factor precursor"
          /codon_start=1
          /protein_id="AAA52787.1"
          /db_xref="GI:306927"
          /db_xref="GDB:G00-120-075"
          /translation="MGKISSLPTQLFKCCFCDFLKVKMHTMSSSHLFYLALCLLTFTS
          SATAGPETLCGAELVDALQFVCGDRGFYFNKPTGYGSSSRAPQTGIVDECCFRSCDL
          RRLEMYCAPLKPASARSVRAQRHTDMPKTQKEVHLKNASRGSAAGNKNYRM"
sig_peptide 149. .292
          /gene="IGFBP1"
          /note="insulin-like growth factor signal peptide"
mat_peptide 293. .502
          /gene="IGFBP1"
          /product="insulin-like growth factor"
BASE COUNT      283 a      251 c      239 g      303 t
ORIGIN          Chromosome 7p13-p12.

```

```

Query Match          66.3%; Score 342.6; DB 9; Length 1076;
Best Local Similarity 87.1%; Pred. No. 8.8e-99;
Matches 454; Conservative 0; Mismatches 14; Indels 53; Gaps 5;

```

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 60
      |||
Db      293 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 352
      |||
Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db      353 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 412
      |||
Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db      413 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 472
      |||
Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      |||
Db      473 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 532
      |||
Qy      241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db      533 ATGCCCAAGACCCAG----- 547
      |||
Qy      301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db      548 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 603
      |||
Qy      361 GATGTA-GAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTGCTC 419
      |||
Db      604 GATGTAGGAAGACCCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 663
      |||

```

Search completed: December 13, 2003, 09:27:32  
Job time : 2313.97 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 02:35:18 ; Search time 207.586 Seconds  
(without alignments)  
6723.048 Million cell updates/sec

Title: US-09-852-261-1  
Perfect score: 517  
Sequence: 1 ggaccggagacgctctgcgg.....tgaaatacacaagtaaacaat 517

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:\*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result	%					
No.	Score	Query Match	Length	DB	ID	Description
1	517	100.0	517	22	AAD06398	Human IGF-I isoform
2	517	100.0	517	24	AAS16877	Human mechano-grow
3	467.4	90.4	523	22	AAD06400	Rabbit IGF-I isofo
4	467.4	90.4	523	24	AAS16879	Rabbit mechano-gro
5	467.4	90.4	553	18	AAT84893	Rabbit insulin lik
6	377.2	73.0	471	22	AAD06405	Rabbit liver-type
7	377.2	73.0	471	24	AAS16884	Rabbit insulin-lik
8	344.2	66.6	818	8	AAN70436	Sequence encoding
9	344.2	66.6	7260	24	ABT11091	Human breast cance
10	344.2	66.6	7260	24	ABK84583	Human cDNA differe
11	344.2	66.6	7260	24	ABN97244	Gene #3742 used to
12	344.2	66.6	7260	24	ABK64812	Human benign prost
13	344.2	66.6	7260	24	ABK35504	Human endometrial
14	344.2	66.6	7260	24	ABK35561	Gene IGF1 differen
15	342.6	66.3	777	18	AAT84894	Human insulin like
16	339.4	65.6	622	7	AAN60490	Human prepro-somat
17	325.2	62.9	539	22	AAD06399	Rat IGF-I isoform
18	325.2	62.9	539	24	AAS16878	Rat mechano-growth
19	318.2	61.5	651	25	ABV76185	Mouse insulin-like
20	308.6	59.7	1136	8	AAN70435	Sequence encoding
21	286.4	55.4	3599	19	AAV50428	Plasmid pIG0552 lo
22	286.4	55.4	3599	19	AAV40796	Actual sequence of
23	286.4	55.4	3600	19	AAV50427	Plasmid pIG0552 up
24	286.4	55.4	3600	19	AAV40795	Expected sequence
25	286.4	55.4	5707	20	AAX88055	Plasmid pIG0335 DN
26	286.4	55.4	6345	20	AAX88054	Plasmid pIG0100A D
27	285.4	55.2	612	22	AAS14695	Human cDNA encodin
28	285.4	55.2	612	25	ABZ83309	Toxicologically re
29	271.2	52.5	978	14	AAQ47804	Sequence encoding
30	258.4	50.0	317	24	AAS16882	Human insulin-like
31	258.4	50.0	318	22	AAD06403	Human liver-type I
32	258.4	50.0	462	19	AAV50426	Human IGF-1 encodi
33	258.4	50.0	462	19	AAV40794	Human IGF-I coding
34	258.4	50.0	462	24	ABZ35734	Human IGF1 polynuc
35	258.4	50.0	462	24	ABX09977	Human IGF1 DNA fra
36	258.4	50.0	462	24	ABV78158	Human IGF1 DNA SEQ
37	258.4	50.0	462	24	ABL91699	Human polynucleoti
38	252.6	48.9	1052	20	AAX27498	Rat liver form of
39	247.8	47.9	487	22	AAD06404	Rat liver-type IGF
40	247.8	47.9	487	24	AAS16883	Rat insulin-like g
41	234.2	45.3	671	24	ABT09479	Phase-1 Rat CT gen
42	210	40.6	210	24	AAD45568	Human insulin-like
43	210	40.6	210	24	AAD44955	Human insulin grow
44	210	40.6	210	24	ABA03146	Native mature IGF-
45	208.4	40.3	237	12	AAQ13568	Beta-gal/IGF-1 fus

# ALIGNMENTS

RESULT 1

AAD06398

ID AAD06398 standard; cDNA; 517 BP.

XX

AC AAD06398;

XX

DT 10-AUG-2001 (first entry)

XX

DE Human IGF-I isoform mechano-growth factor (MGF) cDNA.

XX

KW Human; IGF-I isoform; Insulin-like Growth Factor-I; MGF;  
 KW mechano-growth factor; neurological disorder; neurodegenerative disorder;  
 KW amyotrophic lateral sclerosis; spinal muscular atrophy; muscular atrophy;  
 KW poliomyelitis; post-polio syndrome; toxin; motoneurone disorder;  
 KW nerve damage; autosomal muscular dystrophy; diabetic neuropathy;  
 KW sex-linked muscular dystrophy; peripheral neuropathy;  
 KW Alzheimer's disease; Parkinson's disease; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..333

FT /\*tag= a

FT /product= "Mechano-growth factor (MGF)"

FT /note= "This region comprises exons 3-6. The CDS does  
 FT not include start codon"

FT /partial

XX

PN WO200136483-A1.

XX

PD 25-MAY-2001.

XX

PF 15-NOV-2000; 2000WO-GB04354.

XX

PR 15-NOV-1999; 99GB-0026968.

XX

PA (UNLO ) UNIV COLLEGE LONDON.

XX

PI Goldspink G, Johnson I;

XX

DR WPI; 2001-355620/37.

DR P-PSDB; AAE02447.

XX

PT Use of mechano-growth factor, an isoform of Insulin-like Growth  
 PT Factor-I, capable of reducing motoneurone loss, in the manufacture of a  
 PT medicament for the treatment of neurological disorder -

XX

PS Claim 4; Page 49-50; 66pp; English.

XX

CC The present invention relates to use of mechano-growth factor (MGF),  
 CC an Insulin-like Growth Factor-I (IGF-I) isoform in the manufacture of a  
 CC medicament for the treatment of neurological disorder. The MGF is capable  
 CC of reducing motoneurone loss by 20% or greater in response to nerve  
 CC avulsion, and effects motoneurone rescue, preferably adult motoneurone  
 CC rescue. The MGF polynucleotide and polypeptide are useful in the  
 CC manufacture of a medicament for the treatment of a neurological disorder,



CC including a disorder of motoneurons and/or neurodegenerative disorder,  
 CC e.g., amyotrophic lateral sclerosis, spinal muscular atrophy, progressive  
 CC spinal muscular atrophy, infantile or juvenile muscular atrophy,  
 CC poliomyelitis or post-polio syndrome, a disorder caused by exposure to a  
 CC toxin, motoneuron trauma, a motoneuron lesion or nerve damage, an  
 CC injury that affects motoneurons, motoneuron loss associated with aging,  
 CC autosomal or sex-linked muscular dystrophy, diabetic neuropathy,  
 CC peripheral neuropathies, Alzheimer's disease and Parkinson's disease.  
 CC The present sequence is human IGF-I isoform MGF cDNA. MGF is a muscle  
 CC isoform having extracellular (Ec) domain, hence also referred as  
 CC IGF-I-Ec. The MGF protein comprises amino acid sequences encoded by  
 CC nucleic acid sequence of IGF-I exons 4, 5 and 6 in the reading frame  
 CC of MGF.

XX

SQ Sequence 517 BP; 150 A; 130 C; 139 G; 98 T; 0 other;

Query Match 100.0%; Score 517; DB 22; Length 517;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-146;  
 Matches 517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC	60
Db	1	GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC	60
Qy	61	AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG	120
Db	61	AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG	120
Qy	121	ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT	180
Db	121	ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT	180
Qy	181	TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC	240
Db	181	TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC	240
Qy	241	ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA	300
Db	241	ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA	300
Qy	301	AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG	360
Db	301	AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG	360
Qy	361	GATGTAGAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTGCTCT	420
Db	361	GATGTAGAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTGCTCT	420
Qy	421	GCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA	480
Db	421	GCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA	480
Qy	481	GATGGCATTTCCCCCAATGAAATACACAAGTAAACAT	517
Db	481	GATGGCATTTCCCCCAATGAAATACACAAGTAAACAT	517

RESULT 2

AAS16877

ID AAS16877 standard; cDNA; 517 BP.

XX

AC AAS16877;

XX

DT 25-FEB-2002 (first entry)

XX

DE Human mechano-growth factor (MGF) cDNA.

XX

KW Human; mechano-growth factor; insulin-like growth factor I; IGF-I; MGF;  
 KW neuroprotective; nerve damage; peripheral nervous system; nerve severing;  
 KW muscle; neurological disorder; motoneuron loss; motorneuron disorder; ss;  
 KW nerve avulsion.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS

1..333

FT

/\*tag= a

FT

/product= "Human MGF"

FT

/partial

FT

/note= "No start codon"

FT

exon

1..76

FT

/\*tag= b

FT

/number= 3

FT

exon

77..259

FT

/\*tag= c

FT

/number= 4

FT

exon

260..307

FT

/\*tag= d

FT

/number= 5

FT

exon

308..330

FT

/\*tag= e

FT

/number= 6

XX

PN WO200185781-A2.

XX

PD 15-NOV-2001.

XX

PF 10-MAY-2001; 2001WO-GB02054.

XX

PR 10-MAY-2000; 2000GB-0011278.

XX

PA (UNLO ) UNIV COLLEGE LONDON.

PA

(EGRI-) EAST GRINSTEAD MEDICAL RES TRUST.

XX

PI Goldspink G, Terenghi G;

XX

DR WPI; 2002-055585/07.

DR

P-PSDB; AAU10559.

XX

PT Use of insulin-like growth factor I (IGF-I) isoform known as  
 PT mechano-growth factor which is encoded by IGF-I exons 4,5,6 and has  
 PT ability to reduce motoneuron loss in response to nerve avulsion, to  
 PT treat nerve damage -

XX

PS Claim 11; Fig 5; 65pp; English.

XX

CC The invention relates to the use of an insulin-like growth factor I  
CC (IGF-I) isoform, known as mechano-growth factor (MGF), in the manufacture  
CC of a medicament for treating nerve damage in the peripheral nervous  
CC system, or for treating nerve damage by localising MGF at the site of  
CC damage. The nerve damage may include severing of a nerve. The treatment  
CC may be combined with another treatment (such as a polypeptide growth  
CC factor other than MGF) that prevents or diminishes degeneration of the  
CC target organ (for example, muscle) which the damaged nerve innervates,  
CC whereby the treatment of the muscle with MGF or a polynucleotide encoding  
CC MGF prevents or diminishes degeneration. The method is useful for  
CC treating neurological disorders, preferably motorneuron disorders. These  
CC methods can reduce motoneuron loss by 20% or greater in response to nerve  
CC avulsion. This sequence represents cDNA encoding the human MGF.

XX

SQ Sequence 517 BP; 150 A; 130 C; 139 G; 98 T; 0 other;

Query Match 100.0%; Score 517; DB 24; Length 517;  
Best Local Similarity 100.0%; Pred. No. 1.6e-146;  
Matches 517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC	60
Db	1	GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC	60
Qy	61	AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG	120
Db	61	AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG	120
Qy	121	ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT	180
Db	121	ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT	180
Qy	181	TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC	240
Db	181	TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC	240
Qy	241	ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA	300
Db	241	ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA	300
Qy	301	AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG	360
Db	301	AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG	360
Qy	361	GATGTAGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTCT	420
Db	361	GATGTAGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTCT	420
Qy	421	GCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA	480
Db	421	GCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA	480
Qy	481	GATGGCATTTCCCCCAATGAAATACACAAGTAAACAT	517
Db	481	GATGGCATTTCCCCCAATGAAATACACAAGTAAACAT	517

RESULT 3

AAD06400

ID AAD06400 standard; cDNA; 523 BP.

XX

AC AAD06400;

XX

DT 10-AUG-2001 (first entry)

XX

DE Rabbit IGF-I isoform mechano-growth factor (MGF) cDNA.

XX

KW Rabbit; IGF-I isoform; Insulin-like Growth Factor-I; MGF;  
 KW mechano-growth factor; neurological disorder; neurodegenerative disorder;  
 KW amyotrophic lateral sclerosis; spinal muscular atrophy; muscular atrophy;  
 KW poliomyelitis; post-polio syndrome; toxin; motoneurone disorder;  
 KW nerve damage; autosomal muscular dystrophy; diabetic neuropathy;  
 KW sex-linked muscular dystrophy; peripheral neuropathy;  
 KW Alzheimer's disease; Parkinson's disease; ss.

XX

OS Oryctolagus cuniculus.

XX

FH Key Location/Qualifiers

FT CDS

1..336

FT

/\*tag= a

FT

/product= "Mechano-growth factor (MGF)"

FT

/note= "This region comprises exons 3-6. The CDS does  
 not include start codon"

FT

/partial

FT

XX

XX

PN WO200136483-A1.

XX

PD 25-MAY-2001.

XX

PF 15-NOV-2000; 2000WO-GB04354.

XX

PR 15-NOV-1999; 99GB-0026968.

XX

PA (UNLO ) UNIV COLLEGE LONDON.

XX

PI Goldspink G, Johnson I;

XX

DR WPI; 2001-355620/37.

DR

P-PSDB; AAE02449.

XX

PT Use of mechano-growth factor, an isoform of Insulin-like Growth  
 PT Factor-I, capable of reducing motoneurone loss, in the manufacture of a  
 PT medicament for the treatment of neurological disorder -

XX

PS Claim 4; Page 53-54; 66pp; English.

XX

CC The present invention relates to use of mechano-growth factor (MGF),  
 CC an Insulin-like Growth Factor-I (IGF-I) isoform in the manufacture of a  
 CC medicament for the treatment of neurological disorder. The MGF is capable  
 CC of reducing motoneurone loss by 20% or greater in response to nerve  
 CC avulsion, and effects motoneurone rescue, preferably adult motoneurone  
 CC rescue. The MGF polynucleotide and polypeptide are useful in the

CC manufacture of a medicament for the treatment of a neurological disorder,  
 CC including a disorder of motoneurons and/or neurodegenerative disorder,  
 CC e.g., amyotrophic lateral sclerosis, spinal muscular atrophy, progressive  
 CC spinal muscular atrophy, infantile or juvenile muscular atrophy,  
 CC poliomyelitis or post-polio syndrome, a disorder caused by exposure to a  
 CC toxin, motoneuron trauma, a motoneuron lesion or nerve damage, an  
 CC injury that affects motoneurons, motoneuron loss associated with aging,  
 CC autosomal or sex-linked muscular dystrophy, diabetic neuropathy,  
 CC peripheral neuropathies, Alzheimer's disease and Parkinson's disease.  
 CC The present sequence is rabbit IGF-I isoform MGF cDNA. MGF is a muscle  
 CC isoform having extracellular (Ec) domain, hence also referred as  
 CC IGF-I-Ec. The MGF protein comprises amino acid sequences encoded by  
 CC nucleic acid sequence of IGF-I exons 4, 5 and 6 in the reading frame  
 CC of MGF.

XX

SQ Sequence 523 BP; 154 A; 129 C; 142 G; 98 T; 0 other;

Query Match 90.4%; Score 467.4; DB 22; Length 523;  
 Best Local Similarity 96.2%; Pred. No. 1.8e-131;  
 Matches 501; Conservative 0; Mismatches 16; Indels 4; Gaps 2;

Qy	1	GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC	60
Db	1	GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC	60
Qy	61	AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG	120
Db	61	AGGGGCTTTTATTTCAACAAGCCACAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG	120
Qy	121	ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT	180
Db	121	ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC	180
Qy	181	TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC	240
Db	181	TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCCAGCGCCACACCGAC	240
Qy	241	ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCA--G	297
Db	241	ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAATGAAGTCTCAGAGG	300
Qy	298	AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA	357
Db	301	AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA	360
Qy	358	CAGGATGTA-GAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG	416
Db	361	CAGGATGTAGGAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG	420
Qy	417	CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT	476
Db	421	CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT	480
Qy	477	CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT	517
Db	481	CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT	521

RESULT 4  
 AAS16879  
 ID AAS16879 standard; cDNA; 523 BP.  
 XX  
 AC AAS16879;  
 XX  
 DT 25-FEB-2002 (first entry)  
 XX  
 DE Rabbit mechano-growth factor (MGF) cDNA.  
 XX  
 KW Rabbit; mechano-growth factor; insulin-like growth factor I; IGF-I; MGF;  
 KW neuroprotective; nerve damage; peripheral nervous system; nerve severing;  
 KW muscle; neurological disorder; motoneuron loss; motoneuron disorder; ss;  
 KW nerve avulsion.  
 XX  
 OS Oryctolagus cuniculus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..336  
 FT /\*tag= a  
 FT /product= "Rabbit MGF"  
 FT /partial  
 FT /note= "No start codon"  
 FT exon 1..76  
 FT /\*tag= b  
 FT /number= 3  
 FT exon 77..259  
 FT /\*tag= c  
 FT /number= 4  
 FT exon 260..309  
 FT /\*tag= d  
 FT /number= 5  
 FT exon 311..333  
 FT /\*tag= e  
 FT /number= 6  
 XX  
 PN WO200185781-A2.  
 XX  
 PD 15-NOV-2001.  
 XX  
 PF 10-MAY-2001; 2001WO-GB02054.  
 XX  
 PR 10-MAY-2000; 2000GB-0011278.  
 XX  
 PA (UNLO ) UNIV COLLEGE LONDON.  
 PA (EGRI-) EAST GRINSTEAD MEDICAL RES TRUST.  
 XX  
 PI Goldspink G, Terenghi G;  
 XX  
 DR WPI; 2002-055585/07.  
 DR P-PSDB; AAU10561.  
 XX  
 PT Use of insulin-like growth factor I (IGF-I) isoform known as  
 PT mechano-growth factor which is encoded by IGF-I exons 4,5,6 and has  
 PT ability to reduce motoneuron loss in response to nerve avulsion, to  
 PT treat nerve damage -

XX  
PS Disclosure; Fig 7; 65pp; English.  
XX

CC The invention relates to the use of an insulin-like growth factor I  
CC (IGF-I) isoform, known as mechano-growth factor (MGF), in the manufacture  
CC of a medicament for treating nerve damage in the peripheral nervous  
CC system, or for treating nerve damage by localising MGF at the site of  
CC damage. The nerve damage may include severing of a nerve. The treatment  
CC may be combined with another treatment (such as a polypeptide growth  
CC factor other than MGF) that prevents or diminishes degeneration of the  
CC target organ (for example, muscle) which the damaged nerve innervates,  
CC whereby the treatment of the muscle with MGF or a polynucleotide encoding  
CC MGF prevents or diminishes degeneration. The method is useful for  
CC treating neurological disorders, preferably motorneuron disorders. These  
CC methods can reduce motoneuron loss by 20% or greater in response to nerve  
CC avulsion. This sequence represents cDNA encoding the rabbit MGF.

XX  
SQ Sequence 523 BP; 154 A; 129 C; 142 G; 98 T; 0 other;

Query Match 90.4%; Score 467.4; DB 24; Length 523;  
Best Local Similarity 96.2%; Pred. No. 1.8e-131;  
Matches 501; Conservative 0; Mismatches 16; Indels 4; Gaps 2;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db      1 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

Qy     61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db     61 AGGGGCTTTTATTTCAACAAGCCACAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG 120

Qy    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy    181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      |||
Db    181 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCCAGCGCCACACCGAC 240

Qy    241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCA---G 297
      |||
Db    241 ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAATGAAGTCTCAGAGG 300

Qy    298 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 357
      |||
Db    301 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 360

Qy    358 CAGGATGTA-GAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 416
      |||
Db    361 CAGGATGTAGGAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 420

Qy    417 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT 476
      |||
Db    421 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT 480

Qy    477 CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 517
      |||
```

Db 481 CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 521

RESULT 5

AAT84893

ID AAT84893 standard; cDNA; 553 BP.

XX

AC AAT84893;

XX

DT 14-APR-1998 (first entry)

XX

DE Rabbit insulin like growth factor 1 encoding cDNA.

XX

KW Insulin like growth factor 1; IGF-1; Ec peptide; muscle disorder;

KW heart; neuromuscular disease; primer; ss.

XX

OS Oryctolagus cuniculus.

XX

FH Key Location/Qualifiers

FT CDS 1..366

FT /\*tag= a

FT /product= "IGF-1"

XX

PN WO9733997-A1.

XX

PD 18-SEP-1997.

XX

PF 11-MAR-1997; 97WO-GB00658.

XX

PR 11-MAR-1996; 96GB-0005124.

XX

PA (UNLO ) ROYAL FREE HOSPITAL SCHOOL MED.

XX

PI Goldspink G;

XX

DR WPI; 1997-470877/43.

DR P-PSDB; AAW23301.

XX

PT Use of insulin like growth factor I characterised by presence of Ec

PT peptide - to treat humans or animals, particularly muscle disorders,

PT heart conditions or neuromuscular diseases

XX

PS Disclosure; Fig 3; 33pp; English.

XX

CC A use of insulin like growth factor I (IGF-1) has been developed, and  
CC is characterised by the presence of the Ec peptide, or a functional  
CC equivalent, in the treatment or therapy of a human or animal. The IGF-1  
CC polypeptide can be used to treat muscular disorders, e.g. Duchenne or  
CC Becker muscular dystrophy, autosomal dystrophies and related progressive  
CC skeletal muscle weakness and wasting, muscle atrophy in ageing humans,  
CC spinal cord injury induced muscle atrophy and neuromuscular diseases,  
CC and cardiac disorders, e.g. diseases where promotion of cardiac muscle  
CC protein synthesis is a beneficial treatment, cardiomyopathies and acute  
CC heart failure or insult, specifically myocarditis or myocardial  
CC infarction. It can also be used to promote bone fracture healing and  
CC maintenance of bone in old age. The present sequence encodes rabbit  
CC IGF-1 used in the present specification.



XX

SQ Sequence 553 BP; 159 A; 142 C; 147 G; 105 T; 0 other;

Query Match 90.4%; Score 467.4; DB 18; Length 553;

Best Local Similarity 96.2%; Pred. No. 1.8e-131;

Matches 501; Conservative 0; Mismatches 16; Indels 4; Gaps 2;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db      31 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 90
      |||
Qy     61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db     91 AGGGGCTTTTATTTCAACAAGCCACAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG 150
      |||
Qy    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    151 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 210
      |||
Qy    181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240
      |||
Db    211 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGTCCGTCCGTGCCAGCGCCACACCGAC 270
      |||
Qy    241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCA---G 297
      |||
Db    271 ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAATGAAGTCTCAGAGG 330
      |||
Qy    298 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 357
      |||
Db    331 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 390
      |||
Qy    358 CAGGATGTA-GAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTG 416
      |||
Db    391 CAGGATGTAGGAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTG 450
      |||
Qy    417 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT 476
      |||
Db    451 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT 510
      |||
Qy    477 CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 517
      |||
Db    511 CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 551
```

RESULT 6

AAD06405

ID AAD06405 standard; cDNA; 471 BP.

XX

AC AAD06405;

XX

DT 10-AUG-2001 (first entry)

XX

DE Rabbit liver-type IGF-I isoform (L.IGF-I) cDNA.

XX

KW Rabbit; IGF-I isoform; Insulin-like Growth Factor-I; MGF;

KW mechano-growth factor; neurological disorder; neurodegenerative disorder;

KW amyotrophic lateral sclerosis; spinal muscular atrophy; muscular atrophy;

KW poliomyelitis; post-polio syndrome; toxin; motoneurone disorder;  
 KW nerve damage; autosomal muscular dystrophy; diabetic neuropathy;  
 KW sex-linked muscular dystrophy; peripheral neuropathy;  
 KW Alzheimer's disease; Parkinson's disease; liver; L.IGF-I; ss.  
 XX  
 OS Oryctolagus cuniculus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..318  
 FT /\*tag= a  
 FT /product= "Liver-type IGF-I isoform (L.IGF-I)"  
 FT /transl\_except= (pos:7..9, aa:Gln)  
 FT /transl\_except= (pos:25..27, aa:Gln)  
 FT /note= "These translation exceptions occur while decoding  
 FT the alternative version of the protein (AAE02456).  
 FT The CDS comprises exons 3, 4 and 6 and  
 FT does not include start codon"  
 FT /partial  
 XX  
 PN WO200136483-A1.  
 XX  
 PD 25-MAY-2001.  
 XX  
 PF 15-NOV-2000; 2000WO-GB04354.  
 XX  
 PR 15-NOV-1999; 99GB-0026968.  
 XX  
 PA (UNLO ) UNIV COLLEGE LONDON.  
 XX  
 PI Goldspink G, Johnson I;  
 XX  
 DR WPI; 2001-355620/37.  
 DR P-PSDB; AAE02452, AAE02456.  
 XX  
 PT Use of mechano-growth factor, an isoform of Insulin-like Growth  
 PT Factor-I, capable of reducing motoneurone loss, in the manufacture of a  
 PT medicament for the treatment of neurological disorder -  
 XX  
 PS Disclosure; Page 59-60; 66pp; English.  
 XX  
 CC The present invention relates to use of mechano-growth factor (MGF),  
 CC an Insulin-like Growth Factor-I (IGF-I) isoform in the manufacture of a  
 CC medicament for the treatment of neurological disorder. The MGF is capable  
 CC of reducing motoneurone loss by 20% or greater in response to nerve  
 CC avulsion, and effects motoneurone rescue, preferably adult motoneurone  
 CC rescue. The MGF polynucleotide and polypeptide are useful in the  
 CC manufacture of a medicament for the treatment of a neurological disorder,  
 CC including a disorder of motoneurons and/or neurodegenerative disorder,  
 CC e.g., amyotrophic lateral sclerosis, spinal muscular atrophy, progressive  
 CC spinal muscular atrophy, infantile or juvenile muscular atrophy,  
 CC poliomyelitis or post-polio syndrome, a disorder caused by exposure to a  
 CC toxin, motoneurone trauma, a motoneurone lesion or nerve damage, an  
 CC injury that affects motoneurons, motoneurone loss associated with aging,  
 CC autosomal or sex-linked muscular dystrophy, diabetic neuropathy,  
 CC peripheral neuropathies, Alzheimer's disease and Parkinson's disease.  
 CC The present sequence is rabbit liver-type IGF-I isoform (L.IGF-I) cDNA.  
 CC The L.IGF-I protein comprises amino acid sequences encoded by

CC nucleic acid sequence of IGF-I exons 4 and 6.

XX

SQ Sequence 471 BP; 132 A; 118 C; 131 G; 90 T; 0 other;

Query Match 73.0%; Score 377.2; DB 22; Length 471;  
Best Local Similarity 87.8%; Pred. No. 3.8e-104;  
Matches 455; Conservative 0; Mismatches 13; Indels 50; Gaps 2;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db      1 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

Qy     61 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db     61 AGGGGCTTTTATTTCAACAAGCCCACAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG 120

Qy    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy    181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      |||
Db    181 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCCAGCGCCACACCGAC 240

Qy    241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db    241 ATGCCCAAGACTCAG----- 255

Qy    301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db    256 ----AAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 311

Qy    361 GATGTA-GAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTGCTC 419
      |||
Db    312 GATGTAGGAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTGCTC 371

Qy    420 TGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA 479
      |||
Db    372 TGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA 431

Qy    480 AGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 517
      |||
Db    432 AGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 469
```

RESULT 7

AAS16884

ID AAS16884 standard; cDNA; 471 BP.

XX

AC AAS16884;

XX

DT 25-FEB-2002 (first entry)

XX

DE Rabbit insulin-like growth factor I liver-type isoform (L.IGF-I) cDNA.

XX

KW Rabbit; mechano-growth factor; insulin-like growth factor I; IGF-I; MGF;  
KW neuroprotective; nerve damage; peripheral nervous system; nerve severing;

KW muscle; neurological disorder; motoneuron loss; motorneuron disorder; ss;  
 KW nerve avulsion; insulin-like growth factor I liver-type isoform; L.IGF-I;  
 XX  
 OS *Oryctolagus cuniculus*.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..318  
 FT /\*tag= a  
 FT /product= "Rabbit L.IGF-I"  
 FT /partial  
 FT /note= "No start codon"  
 FT exon 1..75  
 FT /\*tag= b  
 FT /number= exon 3  
 FT exon 76..258  
 FT /\*tag= c  
 FT /number= exon 4  
 FT exon 259..315  
 FT /\*tag= d  
 FT /number= exon 6  
 XX  
 PN WO200185781-A2.  
 XX  
 PD 15-NOV-2001.  
 XX  
 PF 10-MAY-2001; 2001WO-GB02054.  
 XX  
 PR 10-MAY-2000; 2000GB-0011278.  
 XX  
 PA (UNLO ) UNIV COLLEGE LONDON.  
 PA (EGRI-) EAST GRINSTEAD MEDICAL RES TRUST.  
 XX  
 PI Goldspink G, Terenghi G;  
 XX  
 DR WPI; 2002-055585/07.  
 DR P-PSDB; AAU10564.  
 XX  
 PT Use of insulin-like growth factor I (IGF-I) isoform known as  
 PT mechano-growth factor which is encoded by IGF-I exons 4,5,6 and has  
 PT ability to reduce motoneuron loss in response to nerve avulsion, to  
 PT treat nerve damage -  
 XX  
 PS Disclosure; Fig 10; 65pp; English.  
 XX  
 CC The invention relates to the use of an insulin-like growth factor I  
 CC (IGF-I) isoform, known as mechano-growth factor (MGF), in the manufacture  
 CC of a medicament for treating nerve damage in the peripheral nervous  
 CC system, or for treating nerve damage by localising MGF at the site of  
 CC damage. The nerve damage may include severing of a nerve. The treatment  
 CC may be combined with another treatment (such as a polypeptide growth  
 CC factor other than MGF) that prevents or diminishes degeneration of the  
 CC target organ (for example, muscle) which the damaged nerve innervates,  
 CC whereby the treatment of the muscle with MGF or a polynucleotide encoding  
 CC MGF prevents or diminishes degeneration. The method is useful for  
 CC treating neurological disorders, preferably motorneuron disorders. These  
 CC methods can reduce motoneuron loss by 20% or greater in response to nerve  
 CC avulsion. This sequence represents cDNA encoding the rabbit insulin-like

CC growth factor I liver-type isoform (L.IGF-I) used in experiments on  
CC motoneuron loss.

XX

SQ Sequence 471 BP; 132 A; 118 C; 131 G; 90 T; 0 other;

Query Match 73.0%; Score 377.2; DB 24; Length 471;  
Best Local Similarity 87.8%; Pred. No. 3.8e-104;  
Matches 455; Conservative 0; Mismatches 13; Indels 50; Gaps 2;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db      1 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

Qy     61 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db     61 AGGGGCTTTTATTTCAACAAGCCCACAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG 120

Qy    121 ACAGGCATCGTGGATGAGTGCTTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    121 ACAGGCATCGTGGATGAGTGCTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy    181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      |||
Db    181 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGTCCGTCCGTGCCCAGCGCCACACCGAC 240

Qy    241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db    241 ATGCCCAAGACTCAG----- 255

Qy    301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db    256 ----AAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 311

Qy    361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db    312 GATGTAGGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 371

Qy    420 TGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA 479
      |||
Db    372 TGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA 431

Qy    480 AGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 517
      |||
Db    432 AGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 469
```

# RESULT 8

AAN70436

ID AAN70436 standard; cDNA; 818 BP.

XX

AC AAN70436;

XX

DT 25-MAR-2003 (updated)

DT 05-APR-1991 (first entry)

XX

DE Sequence encoding insulin-like growth factor 1A (IGF-1A).

XX

KW Growth promoter; lactation enhancer; cell proliferation; ss.

XX

OS Homo sapiens.

XX

PN EP229750-A.

XX

PD 22-JUL-1987.

XX

PF 06-JAN-1987; 87EP-0870001.

XX

PR 20-NOV-1986; 86US-0929671.

PR 07-JAN-1986; 86US-0816662.

XX

PA (UNIW ) UNIV WASHINGTON.

XX

PI Krivi GG, Rotwein PS;

XX

DR WPI; 1987-200203/29.

XX

PT New pre-pro-insulin-like growth factor-1 protein - obt'd. by

PT recombinant DNA procedures for use as growth promoters for

PT enhancing lactation, for stimulating cell proliferation etc.

XX

PS Example; Fig 5; 59pp; English.

XX

CC A 42 base oligonucleotide corresponding to the DNA sequence encoding  
CC amino acids 10 to 23 of mature human IGF-I was synthesized (AAN70437).

CC The radiolabeled 42 mer was then employed to screen for IGF-I

CC containing DNA sequences in a human liver cDNA library. Insulin-

CC like growth factors-1A and -1B cDNAs were isolated from a human cDNA

CC library by using lambda gt 11 (AAN70435, AAN70436). The human IGF-1

CC genomic gene was isolated and mapped. It encodes at least two

CC preproinsulin-like growth factor-1 proteins. An essentially pure

CC preproinsulin-like growth factor-1 protein comprising the sequence

CC of amino acids shown in Figure six is claimed (AAP70277).

CC (Updated on 25-MAR-2003 to correct PA field.)

XX

SQ Sequence 818 BP; 232 A; 186 C; 187 G; 213 T; 0 other;

Query Match 66.6%; Score 344.2; DB 8; Length 818;

Best Local Similarity 87.3%; Pred. No. 4.8e-94;

Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;

Qy 1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60  
|||  
Db 203 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 262  
Qy 61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120  
|||  
Db 263 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 322  
Qy 121 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180  
|||  
Db 323 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 382  
Qy 181 TGCGCACCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240  
|||

Db 383 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 442  
 QY 241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300  
 |||||  
 Db 443 ATGCCCAAGACCCAG----- 457  
 QY 301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360  
 |||||  
 Db 458 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 513  
 QY 361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419  
 |||||  
 Db 514 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 573  
 QY 420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTC 477  
 |||||  
 Db 574 TGCACGAGTTACCTGTTAAACTTTGGAACACCTACCAAAAAATAAGTTTGATAACATTTA 633  
 QY 478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517  
 |||||  
 Db 634 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 674

RESULT 9

ABT11091

ID ABT11091 standard; cDNA; 7260 BP.

XX

AC ABT11091;

XX

DT 04-DEC-2002 (first entry)

XX

DE Human breast cancer associated coding sequence SEQ ID NO: 1225.

XX

KW Human; breast specific gene; breast cancer; differential expression;  
 KW cytostatic; gene therapy; gene; ss.

XX

OS Homo sapiens.

XX

PN WO200259271-A2.

XX

PD 01-AUG-2002.

XX

PF 25-JAN-2002; 2002WO-US02176.

XX

PR 25-JAN-2001; 2001US-263757P.

PR 25-APR-2001; 2001US-286090P.

PR 23-MAY-2001; 2001US-292517P.

XX

PA (GENE-) GENE LOGIC INC.

XX

PI Orr MS, Nation M, Diggans JC, Zeng W;

XX

DR WPI; 2002-674803/72.

XX

PT Diagnosing breast cancer in a patient comprises detecting the level of  
 PT gene expression in cell or tissue samples, where a differential gene  
 PT expression is indicative of breast cancer -

XX  
PS  
XX  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
XX  
SQ

Claim 1; SEQ ID NO 1225; 260pp + Sequence Listing; English.

The present invention relates to methods of diagnosing breast cancer in a patient, which comprise detecting the level of expression in a tissue sample of two or more genes selected from those shown in ABT09867-ABT11112, where a differential expression of the genes indicates breast cancer. The methods are useful in diagnosing, treating, detecting the progression, and in monitoring treatment of breast cancer in patients. The methods are also useful as a screening tool for agents that modulate the onset or progression of breast cancer. The breast cancer genes may be used as diagnostic markers for the prediction or identification of the malignant state of breast tissue, for confirming the type and progression of cancer, and for drug screening and assays. The present sequence is a coding sequence of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub.published\_pct\_sequences.

Sequence 7260 BP; 2330 A; 1415 C; 1240 G; 2275 T; 0 other;

Query Match 66.6%; Score 344.2; DB 24; Length 7260;  
Best Local Similarity 87.3%; Pred. No. 1.1e-93;  
Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db    311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 370

Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db    371 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 430

Qy     121 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    431 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490

Qy     181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      |||
Db    491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 550

Qy     241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db    551 ATGCCCAAGACCCAG----- 565

Qy     301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db    566 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 621

Qy     361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db    622 GATGTAGGAAGACCCTCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 681

Qy     420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTT 477
      |||
Db    682 TGCACGAGTTACCTGTTAAACTTTGGAACACCTACCAAAAAATAAGTTTGATAACATTTA 741
```



Qy 478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517  
 ||||| ||| ||||||||||||||||||||  
 Db 742 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 782

RESULT 10

ABK84583

ID ABK84583 standard; cDNA; 7260 BP.

XX

AC ABK84583;

XX

DT 14-AUG-2002 (first entry)

XX

DE Human cDNA differentially expressed in granulocytic cells #1154.

XX

KW Human; ss; granulocytic cell; DNA chip; bacterial infection;  
 KW viral infection; parasitic infection; protozoal infection;  
 KW fungal infection; sterile inflammatory disease; psoriasis;  
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;  
 KW adult respiratory distress syndrome; inflammatory bowel disease;  
 KW Crohn's disease; ulcerative colitis; periodontal disease;  
 KW granulocyte activation; chronic inflammation; allergy.

XX

OS Homo sapiens.

XX

PN WO200228999-A2.

XX

PD 11-APR-2002.

XX

PF 03-OCT-2001; 2001WO-US30821.

XX

PR 03-OCT-2000; 2000US-237189P.

XX

PA (GENE-) GENE LOGIC INC.

XX

PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX

DR WPI; 2002-435328/46.

XX

PT Detecting granulocyte activation by detecting differential expression  
 PT of genes associated with granulocyte activation, which serves as  
 PT diagnostic markers that is useful for monitoring disease states and  
 PT drug toxicity -

XX

PS Claim 1; SEQ ID No 1154; 114pp; English.

XX

CC The invention relates to detecting (M1) granulocyte (GC) activation  
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by  
 CC DNA chip analysis as given in the specification, and comparing  
 CC the expression level to an expression level in an unactivated  
 CC GC, where differential expression of Gs is indicative of GCA.  
 CC Also included are modulating (M2) GA by contacting GC with an agent  
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)  
 CC for an agent capable of modulating GCA or an inflammation (especially  
 CC chronic) in a tissue, an allergic response in a subject, exposure of a  
 CC subject to a pathogen or sterile inflammatory disease using the





CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
CC liver carcinoma in a patient. The method is useful for identifying  
CC expression profiles which serve as useful diagnostic markers as well as  
CC markers that can be used to monitor disease states, disease progression,  
CC drug toxicity, drug efficacy and drug metabolism.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 7260 BP; 2330 A; 1415 C; 1240 G; 2275 T; 0 other;

Query Match 66.6%; Score 344.2; DB 24; Length 7260;  
Best Local Similarity 87.3%; Pred. No. 1.1e-93;  
Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db    311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 370

Qy     61 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db    371 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 430

Qy    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    431 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490

Qy    181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      |||
Db    491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 550

Qy    241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db    551 ATGCCCAAGACCCAG----- 565

Qy    301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAAGTACAG 360
      |||
Db    566 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAAGTACAG 621

Qy    361 GATGTA-GAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTGCTC 419
      |||
Db    622 GATGTAGGAAGACCCCTCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 681

Qy    420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCATTTC 477
      |||
Db    682 TGCACGAGTTACCTGTTAAACTTTGGAACACCTACCAAAAAAATAAGTTTGATAACATTTA 741

Qy    478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||
Db    742 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 782
```

RESULT 12

ABK64812

ID ABK64812 standard; DNA; 7260 BP.

XX

AC ABK64812;

XX  
 DT 18-JUN-2002 (first entry)  
 XX  
 DE Human benign prostatic hyperplasia gene #707.  
 XX  
 KW Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200212440-A2.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PF 07-AUG-2001; 2001WO-US24708.  
 XX  
 PR 07-AUG-2000; 2000US-223323P.  
 PR 05-JUN-2001; 2001US-0873319.  
 XX  
 PA (GENE-) GENE LOGIC INC.  
 PA (NISB ) JAPAN TOBACCO INC.  
 XX  
 PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;  
 XX  
 DR WPI; 2002-257476/30.  
 XX  
 PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by  
 PT detecting expression levels of one or more genes in prostate cells from  
 PT patient that are differentially regulated compared to normal prostate  
 PT cells -  
 XX  
 PS Disclosure; Page 391-393; 444pp; English.  
 XX  
 CC The invention relates to a method of diagnosing (I) the onset or  
 CC progression of benign prostatic hyperplasia (BPH), or screening (II) for  
 CC or identifying an agent that modulates the onset or progression of BPH.  
 CC The method is based on changes in gene expression in BPH tissue isolated  
 CC from patients exhibiting different clinical states of prostate  
 CC hyperplasia as compared to normal prostate tissue. (I) comprises  
 CC detecting the expression levels of one or more genes in prostate cells  
 CC from the subject that are differentially regulated compared to normal  
 CC prostate cells. (II) comprises preparing a first gene expression profile  
 CC of BPH cells or BPH-like cell population, exposing the cells to the  
 CC agent, preparing a second gene expression profile of the agent exposed  
 CC cells, and comparing the first and second gene expression profiles.  
 CC (I) is useful for diagnosing the onset or progression of BPH. (II) is  
 CC useful for identifying an agent that modulates the onset or progression  
 CC of BPH. The methods are useful to present information identifying  
 CC the expression level in a tissue or cells, by comparing the expression  
 CC level of genes given in the specification in the tissue or cells to the  
 CC level of expression of gene in the database, and displaying the  
 CC expression levels of at least one gene in the tissue or cell sample  
 CC compared to the expression level in BPH. Agents using (II) are useful for  
 CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human  
 CC benign prostatic hyperplasia gene sequences of the invention.  
 XX  
 SQ Sequence 7260 BP; 2330 A; 1415 C; 1240 G; 2275 T; 0 other;

Query Match 66.6%; Score 344.2; DB 24; Length 7260;  
 Best Local Similarity 87.3%; Pred. No. 1.1e-93;  
 Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db    311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 370

Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db    371 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 430

Qy     121 ACAGGCATCGTGGATGAGTGTCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    431 ACAGGCATCGTGGATGAGTGTCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490

Qy     181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240
      |||
Db    491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 550

Qy     241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db    551 ATGCCCAAGACCCAG----- 565

Qy     301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db    566 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 621

Qy     361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db    622 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 681

Qy     420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTTC 477
      |||
Db    682 TGCACGAGTTACCTGTTAAACTTTGGAACACCTACCAAAAAAATAAGTTTGATAACATTTA 741

Qy     478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||
Db    742 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 782
  
```

# RESULT 13

ABK35504

ID ABK35504 standard; DNA; 7260 BP.

XX

AC ABK35504;

XX

DT 08-MAY-2002 (first entry)

XX

DE Human endometrial cancer related gene, IGF1.

XX

KW Human; ds; gene; endometrial cancer; differential expression;

KW DNA microarray; protein microarray.

XX

OS Homo sapiens.

XX

PN WO200209573-A2.

```
XX PD 07-FEB-2002.
XX PF 31-JUL-2001; 2001WO-US24104.
XX PR 31-JUL-2000; 2000US-221735P.
XX PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX PI Mutter GL;
XX DR WPI; 2002-179967/23.
XX DR P-PSDB; AAU84284.
XX PT Diagnosing endometrial cancer comprises determining expression of
PT nucleic acid molecules or expression products that are differentially
PT expressed in normal and malignant endometrium -
XX PS Claim 1; Page 85-89; 233pp; English.
XX CC The invention relates to diagnosing endometrial cancer in a subject
CC suspected of having endometrial cancer comprising determining the
CC expression of a set of nucleic acid molecules or expression products in
CC an endometrial sample suspected of being cancerous, where the set of
CC nucleic acid molecules comprises at least 2 nucleic acid molecules
CC selected from 50 fully defined sequences as given in the specification.
CC The nucleic acids are used as an array of at least 2 of the 50
CC nucleic acids bound to a solid substrate. Also included is a solid-phase
CC protein microarray comprising at least 2 antibodies or its antigen
CC binding fragments, that specifically bind at least 2 different
CC polypeptides from the 50 fully defined sequences as given in the
CC specification, fixed to a solid substrate. The methods and arrays are
CC useful for the diagnosis of endometrial cancer, selecting and monitoring
CC treatment regimes and identification of lead compounds useful for the
CC treatment of endometrial cancer. The present sequence is one of 50
CC genes differentially expressed between cancerous and non-cancerous
CC samples.
XX SQ Sequence 7260 BP; 2330 A; 1415 C; 1240 G; 2275 T; 0 other;
Query Match          66.6%; Score 344.2; DB 24; Length 7260;
Best Local Similarity 87.3%; Pred. No. 1.1e-93;
Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;
```

```
QY      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 60
        |||||||
Db    311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 370

QY      61 AGGGGCTTTTATTTCACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
        |||||||
Db    371 AGGGGCTTTTATTTCACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 430

QY     121 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
        |||||||
Db   431 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490

QY     181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240
```

```

      |||
Db      491 TGCGCACCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 550
Qy      241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db      551 ATGCCCAAGACCCAG----- 565
Qy      301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db      566 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 621
Qy      361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db      622 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 681
Qy      420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTTC 477
      |||
Db      682 TGCACGAGTTACCTGTAAACTTTGGAACACCTACCAAAAAATAAGTTTGATAACATTTA 741
Qy      478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||
Db      742 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 782

```

#### RESULT 14

ABK35561

ID ABK35561 standard; DNA; 7260 BP.

XX

AC ABK35561;

XX

DT 08-MAY-2002 (first entry)

XX

DE Gene IGF1 differentially expressed in breast cancer tissue.

XX

KW Human; diagnosis of breast cancer; endometrial cancer; breast tumour;

KW MAI; mitotic activity index; cytostatic; gene; ds.

XX

OS Homo sapiens.

XX

PN WO200210436-A2.

XX

PD 07-FEB-2002.

XX

PF 27-JUL-2001; 2001WO-US23642.

XX

PR 28-JUL-2000; 2000US-222093P.

XX

PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.

PA (BAAK/) BAAK J.

XX

PI Baak J, Mutter GL;

XX

DR WPI; 2002-180084/23.

DR P-PSDB; AAU84341.

XX

PT Diagnosing breast cancer comprises determining expression of nucleic

PT acid molecules or expression products that are differentially expressed



PT in normal and malignant tissue -  
XX  
PS Claim 1; Page 74-78; 219pp; English.  
XX

CC The present invention relates to a method for diagnosing breast cancer  
CC in a subject suspected of having endometrial cancer. The method  
CC comprises determining the expression of a set of human genes or  
CC expression products in an endometrial sample suspected of being  
CC cancerous. The human genes of the invention are differentially  
CC expressed in breast tumours characterised as high or low MAI (mitotic  
CC activity index). These sets of genes can be used to discriminate between  
CC high and low MAI breast tumours. The invention also provides DNA and  
CC protein microarrays for analysing the expression of the human genes and  
CC their protein products. The methods and arrays are useful for the  
CC diagnosis and prognosis of endometrial cancer, selecting and monitoring  
CC treatment regimes, and identification of compounds useful for the  
CC treatment of endometrial cancer. ABK35531-ABK35581 represent the human  
CC genes of the invention that are differentially expressed in breast  
CC cancer tissue.

XX  
SQ Sequence 7260 BP; 2330 A; 1415 C; 1240 G; 2275 T; 0 other;

Query Match 66.6%; Score 344.2; DB 24; Length 7260;  
Best Local Similarity 87.3%; Pred. No. 1.1e-93;  
Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db    311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 370

Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db    371 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 430

Qy     121 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    431 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490

Qy     181 TGCGCACCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      |||
Db    491 TGCGCACCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 550

Qy     241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db    551 ATGCCCAAGACCCAG----- 565

Qy     301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db    566 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 621

Qy     361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db    622 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 681

Qy     420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTT 477
      |||
Db    682 TGCACGAGTTACCTGTAAACTTTGGAACACCTACCAAAAAAATAAGTTTGATAACATTTA 741
```

Qy 478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517  
 ||||| ||| ||||||||||||||||||||||||  
 Db 742 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 782

RESULT 15

AAT84894

ID AAT84894 standard; cDNA; 777 BP.

XX

AC AAT84894;

XX

DT 14-APR-1998 (first entry)

XX

DE Human insulin like growth factor 1 Ea isoform encoding cDNA.

XX

KW Insulin like growth factor 1; IGF-1; Ec peptide; muscle disorder;  
 KW heart; neuromuscular disease; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 26..496

FT /\*tag= a

FT /product= "IGF-1 Ea isoform"

XX

PN WO9733997-A1.

XX

PD 18-SEP-1997.

XX

PF 11-MAR-1997; 97WO-GB00658.

XX

PR 11-MAR-1996; 96GB-0005124.

XX

PA (UNLO ) ROYAL FREE HOSPITAL SCHOOL MED.

XX

PI Goldspink G;

XX

DR WPI; 1997-470877/43.

DR P-PSDB; AAW23302.

XX

PT Use of insulin like growth factor I characterised by presence of Ec  
 PT peptide - to treat humans or animals, particularly muscle disorders,  
 PT heart conditions or neuromuscular diseases

XX

PS Disclosure; Fig 4; 33pp; English.

XX

CC A use of insulin like growth factor I (IGF-1) has been developed, and  
 CC is characterised by the presence of the Ec peptide, or a functional  
 CC equivalent, in the treatment or therapy of a human or animal. The IGF-1  
 CC polypeptide can be used to treat muscular disorders, e.g. Duchenne or  
 CC Becker muscular dystrophy, autosomal dystrophies and related progressive  
 CC skeletal muscle weakness and wasting, muscle atrophy in ageing humans,  
 CC spinal cord injury induced muscle atrophy and neuromuscular diseases,  
 CC and cardiac disorders, e.g. diseases where promotion of cardiac muscle  
 CC protein synthesis is a beneficial treatment, cardiomyopathies and acute  
 CC heart failure or insult, specifically myocarditis or myocardial

CC infarction. It can also be used to promote bone fracture healing and  
CC maintenance of bone in old age. The present sequence encodes human  
CC IGF-1 Ea isoform used in the present specification.

XX

SQ Sequence 777 BP; 201 A; 193 C; 204 G; 179 T; 0 other;

Query Match 66.3%; Score 342.6; DB 18; Length 777;  
Best Local Similarity 87.1%; Pred. No. 1.4e-93;  
Matches 454; Conservative 0; Mismatches 14; Indels 53; Gaps 5;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db     179 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 238

Qy     61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db    239 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 298

Qy    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    299 ACAGGTATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 358

Qy    181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240
      |||
Db    359 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 418

Qy    241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db    419 ATGCCCAAGACCCAG----- 433

Qy    301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db    434 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 489

Qy    361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db    490 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 549

Qy    420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTTC 477
      |||
Db    550 TGCACGAGTTACCTGTAAACTTTGGAACACCTACCAAAAAAATAAGTTTGATAACATTTA 609

Qy    478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||
Db    610 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 650
```

Search completed: December 13, 2003, 06:03:48  
Job time : 209.586 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 06:03:55 ; Search time 47.8037 Seconds  
(without alignments)  
4773.589 Million cell updates/sec

Title: US-09-852-261-1  
Perfect score: 517  
Sequence: 1 ggaccggagacgctctgcgg.....tgaaatacacaagtaaacaat 517

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	467.4	90.4	553	3	US-09-142-583A-3 Sequence 3, Appli
2	467.4	90.4	553	3	US-09-142-583A-5 Sequence 5, Appli
3	342.6	66.3	777	3	US-09-142-583A-10 Sequence 10, Appl
4	339.4	65.6	622	6	5405942-2 Patent No. 5405942
5	286.4	55.4	5707	2	US-08-472-809B-8 Sequence 8, Appli
6	286.4	55.4	6345	2	US-08-472-809B-7 Sequence 7, Appli
7	255.2	49.4	357	6	5405942-13 Patent No. 5405942
8	253.6	49.1	357	6	5405942-9 Patent No. 5405942
9	208.4	40.3	210	6	5405942-7 Patent No. 5405942
10	208.4	40.3	210	6	5405942-11 Patent No. 5405942
11	208.4	40.3	2862	4	US-09-255-829-13 Sequence 13, Appl

12	206.8	40.0	210	6	5405942-15	Patent No. 5405942
13	202.8	39.2	240	1	US-08-308-196A-1	Sequence 1, Appli
14	202.8	39.2	240	5	PCT-US91-06452-1	Sequence 1, Appli
15	202.8	39.2	390	3	US-09-029-267-13	Sequence 13, Appl
16	174.4	33.7	798	1	US-07-953-230A-6	Sequence 6, Appli
17	163.4	31.6	770	1	US-07-953-230A-1	Sequence 1, Appli
18	163.4	31.6	846	1	US-07-953-230A-5	Sequence 5, Appli
19	125.8	24.3	485	1	US-07-989-845-29	Sequence 29, Appl
20	125.8	24.3	485	1	US-07-989-844-13	Sequence 13, Appl
21	125.8	24.3	485	1	US-08-110-663-1	Sequence 1, Appli
22	125.8	24.3	485	1	US-08-169-688-1	Sequence 1, Appli
23	125.8	24.3	485	1	US-08-240-121-13	Sequence 13, Appl
24	125.8	24.3	485	1	US-08-451-241-13	Sequence 13, Appl
25	125.8	24.3	485	1	US-08-110-664-1	Sequence 1, Appli
26	125.8	24.3	485	1	US-08-446-882-1	Sequence 1, Appli
27	125.8	24.3	485	1	US-08-385-187A-1	Sequence 1, Appli
28	125.8	24.3	485	1	US-08-470-108-1	Sequence 1, Appli
29	125.8	24.3	485	5	PCT-US93-11297-13	Sequence 13, Appl
30	125.8	24.3	485	5	PCT-US93-11298-29	Sequence 29, Appl
31	123.8	23.9	621	3	US-08-989-251-40	Sequence 40, Appl
32	123.8	23.9	621	3	US-09-340-250-40	Sequence 40, Appl
33	123.8	23.9	621	4	US-09-528-108-40	Sequence 40, Appl
34	122	23.6	237	1	US-07-764-655D-8	Sequence 8, Appli
35	120.8	23.4	243	2	US-08-482-182-75	Sequence 75, Appl
36	120.4	23.3	233	1	US-08-444-142-3	Sequence 3, Appli
37	120.4	23.3	233	1	US-08-444-131-3	Sequence 3, Appli
c 38	120.4	23.3	237	1	US-07-764-655D-9	Sequence 9, Appli
39	120.4	23.3	717	1	US-08-284-784-40	Sequence 40, Appl
40	120.4	23.3	717	2	US-08-854-811-40	Sequence 40, Appl
41	120.4	23.3	783	1	US-08-284-784-43	Sequence 43, Appl
42	120.4	23.3	783	2	US-08-854-811-43	Sequence 43, Appl
43	120.4	23.3	891	1	US-08-284-784-33	Sequence 33, Appl
44	120.4	23.3	891	1	US-08-284-784-34	Sequence 34, Appl
45	120.4	23.3	891	2	US-08-854-811-33	Sequence 33, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-142-583A-3

; Sequence 3, Application US/09142583A

; Patent No. 6221842

; GENERAL INFORMATION:

; APPLICANT: GOLDSPIK, GEOFFREY

; TITLE OF INVENTION: METHOD OF TREATING MUSCULAR DISORDERS

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHYE P.C.

; STREET: 1100 NORTH GLEBE ROAD

; CITY: ARLINGTON

; STATE: VA

; COUNTRY: USA

; ZIP: 22201

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

```

;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.25
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/142,583A
;      FILING DATE: 29-Oct-1998
;      CLASSIFICATION: <Unknown>
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: WO PCT/GB97/00658
;      FILING DATE: 11-MAR-1997
;      APPLICATION NUMBER: GB 9605124.8
;      FILING DATE: 11-MAR-1996
;      ATTORNEY/AGENT INFORMATION:
;      NAME: SADOFF, B. J.
;      REGISTRATION NUMBER: 36663
;      REFERENCE/DOCKET NUMBER: 117-263
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 7038164000
;      TELEFAX: 7038164100
;      INFORMATION FOR SEQ ID NO: 3:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 553 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: both
;      TOPOLOGY: linear
;      MOLECULE TYPE: cDNA
;      FEATURE:
;      NAME/KEY: CDS
;      LOCATION: 1..363
;      SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-142-583A-3

```

```

Query Match          90.4%;  Score 467.4;  DB 3;  Length 553;
Best Local Similarity 96.2%;  Pred. No. 1.2e-134;
Matches 501;  Conservative 0;  Mismatches 16;  Indels 4;  Gaps 2;

```

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
        |||
Db      31 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 90
        |||

Qy      61 AGGGGCTTTTATTTCAACAAGCCCAAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
        |||
Db      91 AGGGGCTTTTATTTCAACAAGCCCAAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG 150
        |||

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
        |||
Db      151 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 210
        |||

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240
        |||
Db      211 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCAGCGCCACACCGAC 270
        |||

Qy      241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCA---G 297
        |||
Db      271 ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAATGAAGTCTCAGAGG 330
        |||

Qy      298 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 357
        |||

```



; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 341..397  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-142-583A-5

Query Match 90.4%; Score 467.4; DB 3; Length 553;  
Best Local Similarity 96.2%; Pred. No. 1.2e-134;  
Matches 501; Conservative 0; Mismatches 16; Indels 4; Gaps 2;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
          |||
Db      31 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 90

Qy     61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
          |||
Db     91 AGGGGCTTTTATTTCAACAAGCCACAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG 150

Qy    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
          |||
Db    151 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 210

Qy    181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
          |||
Db    211 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCCAGCGCCACACCGAC 270

Qy    241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCA---G 297
          |||
Db    271 ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAATGAAGTCTCAGAGG 330

Qy    298 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 357
          |||
Db    331 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 390

Qy    358 CAGGATGTA-GAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 416
          |||
Db    391 CAGGATGTAGGAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 450

Qy    417 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT 476
          |||
Db    451 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTT 510

Qy    477 CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 517
          |||
Db    511 CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 551
```

RESULT 3

US-09-142-583A-10

; Sequence 10, Application US/09142583A

; Patent No. 6221842

; GENERAL INFORMATION:

; APPLICANT: GOLDSPIK, GEOFFREY

; TITLE OF INVENTION: METHOD OF TREATING MUSCULAR DISORDERS

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:



```

;      ADDRESSEE: NIXON & VANDERHYE P.C.
;      STREET: 1100 NORTH GLEBE ROAD
;      CITY: ARLINGTON
;      STATE: VA
;      COUNTRY: USA
;      ZIP: 22201
;
;  COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.25
;
;  CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/142,583A
;      FILING DATE: 29-Oct-1998
;      CLASSIFICATION: <Unknown>
;
;  PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: WO PCT/GB97/00658
;      FILING DATE: 11-MAR-1997
;      APPLICATION NUMBER: GB 9605124.8
;      FILING DATE: 11-MAR-1996
;
;  ATTORNEY/AGENT INFORMATION:
;      NAME: SADOFF, B. J.
;      REGISTRATION NUMBER: 36663
;      REFERENCE/DOCKET NUMBER: 117-263
;
;  TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 7038164000
;      TELEFAX: 7038164100
;
;  INFORMATION FOR SEQ ID NO: 10:
;      SEQUENCE CHARACTERISTICS:
;          LENGTH: 777 base pairs
;          TYPE: nucleic acid
;          STRANDEDNESS: both
;          TOPOLOGY: linear
;
;      MOLECULE TYPE: cDNA
;      FEATURE:
;          NAME/KEY: CDS
;          LOCATION: 26..493
;
;      SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-142-583A-10

```

```

Query Match          66.3%;  Score 342.6;  DB 3;  Length 777;
Best Local Similarity 87.1%;  Pred. No. 4.4e-96;
Matches 454;  Conservative 0;  Mismatches 14;  Indels 53;  Gaps 5;

```

```

Qy      1  GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      179 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 238

Qy      61  AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      239 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 298

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db      299 ACAGGTATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 358

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240

```

```

      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      359 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 418
Qy      241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      ||||||||||||||||
Db      419 ATGCCCAAGACCCAG----- 433
Qy      301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAAGTACAG 360
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      434 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAAGTACAG 489
Qy      361 GATGTA-GAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTGCTC 419
      |||||| |||||||| |||||||||||| ||| |||||||| ||||||||
Db      490 GATGTAGGAAGACCCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 549
Qy      420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACAATTTTC 477
      ||||| |||||||| |||||||| ||| |||||||| |||||||| ||||||||
Db      550 TGCACGAGTTACCTGTTAAACTTTGGAACACCTACCAAAAAAATAAGTTTGATAACATTTA 609
Qy      478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||||| ||| ||||||||||||||||||||||||||||
Db      610 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 650

```

RESULT 4  
5405942-2

```

;Patent No. 5405942
;  APPLICANT: BELL, GRAEME I.;RALL, LESLIE B.;MERRYWEATHER,
;JAMES P.
;  TITLE OF INVENTION: PREPRO INSULIN-LIKE GROWTH FACTORS
;I AND II
;  NUMBER OF SEQUENCES: 16
;  CURRENT APPLICATION DATA:
;    APPLICATION NUMBER: US/07/65,673
;    FILING DATE: 16-JUN-1987
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER: 630,557
;    FILING DATE: 19-JUL-1984
;SEQ ID NO:2:
;  LENGTH: 622
5405942-2

```

```

Query Match          65.6%; Score 339.4; DB 6; Length 622;
Best Local Similarity 69.7%; Pred. No. 3.9e-95;
Matches 363; Conservative 89; Mismatches 16; Indels 53; Gaps 5;

```

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||||||||||||:|:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      45 GGACCGGAGACGCUCUGCGGGGCGAGCUGGUGGAUGCUCUUCAGUUCGUGUGUGGAGAC 104
Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      105 AGGGGCUUUUAUUUCAACAAGCCACAGGGUAUGGCUCAGCAGUCGAGGGGCGCCUCAG 164
Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      ||||| |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      165 ACAGGUAUCGUGGAUGAGUGCUGCUUCCGAGCUGUGAUCAUAGGAGGCUGGAGAUGUAU 224

```

Qy 181 TGCGCACCCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240  
 :|||||||:||||||:||||||:||||:||||:||||:|||||||  
 Db 225 UGCGCACCCCUC AAGCCUGCCAAGUCAGCUCGUCUGUCCGUGCCCAGCGCCACACCGAC 284  
 Qy 241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300  
 |:|||||  
 Db 285 AUGCCCAAGACCCAG----- 299  
 Qy 301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360  
 |||||:||||:||||| ||||:||||||:|||||||:|||||  
 Db 300 ----AAGGAAGUACAUUUGAAGAACGCAAGUAGAGGGAGUGCAGGAAACAAGAACUACAG 355  
 Qy 361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419  
 ||:|:| |||||:|:|||||:||||| |||| |||||:||||:|:|:|  
 Db 356 GAUGUAGGAAGACCCUCCUGAGGAGUGAAGAGUGACAUGCCACCGCAGGAUCCUUUGCUC 415  
 Qy 420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTTC 477  
 :|||| |:|:|:|:| :|||| :|:| ||| |||||:||||:|:|:|:|:|:|  
 Db 416 UGCACGAGUUACCUGUUAAACUUUGGAAACCCUACCAAAAAUAAGUUUGAUAACAUUUA 475  
 Qy 478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517  
 |||||: ||| :|:|:|||||:||||:||||||:|||||:|  
 Db 476 AAAGAUGGGCGUUUCCCCCAUGAAAUACACAAGUAAACAU 516

RESULT 5

US-08-472-809B-8

; Sequence 8, Application US/08472809B

; Patent No. 5925564

; GENERAL INFORMATION:

; APPLICANT: Schwartz, Robert J.

; APPLICANT: DeMayo, Franco J.

; APPLICANT: O'Malley, Bert W.

; TITLE OF INVENTION: Expression Vector Systems and

; TITLE OF INVENTION: Method of Use

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; STREET: Suite 4700

; CITY: Los Angeles

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 90071-2066

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; MEDIUM TYPE: storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/472,809B

; FILING DATE: June 7, 1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/209,846

```

; FILING DATE: March 9, 1994
; APPLICATION NUMBER: 07/789,919
; FILING DATE: No. 5925564ember 6, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 214/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5707 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-472-809B-8

```

```

Query Match          55.4%; Score 286.4; DB 2; Length 5707;
Best Local Similarity 85.6%; Pred. No. 2.3e-78;
Matches 363; Conservative 0; Mismatches 11; Indels 50; Gaps 2;

```

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db      793 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 852
      |||

Qy      61 AGGGGCTTTTATTTCAACAAGCCCAAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db      853 AGGGGCTTTTATTTCAACAAGCCCAAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 912
      |||

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db      913 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 972
      |||

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      |||
Db      973 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 1032
      |||

Qy      241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db      1033 ATGCCCAAGACCCAG----- 1047
      |||

Qy      301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db      1048 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 1103
      |||

Qy      361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db      1104 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCCCCGGGC 1163
      |||

Qy      420 TGCA 423
      |||
Db      1164 TGCA 1167

```

US-08-472-809B-7

US-08-472-809B-7

Matches 363; Conservative 0; Mismatches 11; Indels 50; Gaps 2;

Qy 1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

```

Db      3702 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 3761
Qy      61 AGGGGCTTTTATTTCAACAAGCCCAAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db      3762 AGGGGCTTTTATTTCAACAAGCCCAAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 3821
Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db      3822 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 3881
Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240
      |||
Db      3882 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 3941
Qy      241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db      3942 ATGCCCAAGACCCAG----- 3956
Qy      301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db      3957 ---AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 4012
Qy      361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db      4013 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCCCCGGGC 4072
Qy      420 TGCA 423
      |||
Db      4073 TGCA 4076

```

RESULT 7

5405942-13

;Patent No. 5405942

; APPLICANT: BELL, GRAEME I.; RALL, LESLIE B.; MERRYWEATHER,  
; JAMES P.

; TITLE OF INVENTION: PREPRO INSULIN-LIKE GROWTH FACTORS  
; I AND II

; NUMBER OF SEQUENCES: 16

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/65,673

; FILING DATE: 16-JUN-1987

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 630,557

; FILING DATE: 19-JUL-1984

;SEQ ID NO:13:

; LENGTH: 357

5405942-13

Query Match 49.4%; Score 255.2; DB 6; Length 357;

Best Local Similarity 98.8%; Pred. No. 2.8e-69;

Matches 257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||

```

```

Db      43 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 102

```

```

Qy      61 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db      103 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 162

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db      163 ACAGGTATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 222

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      |||
Db      223 TGCGCACCCCTCAGGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 282

Qy      241 ATGCCCAAGACCCAGAAGTA 260
      |||
Db      283 ATGCCCAAGACCCAGAAGGA 302

```

RESULT 8

5405942-9

;Patent No. 5405942

; APPLICANT: BELL, GRAEME I.; RALL, LESLIE B.; MERRYWEATHER,  
; JAMES P.

; TITLE OF INVENTION: PREPRO INSULIN-LIKE GROWTH FACTORS  
; I AND II

; NUMBER OF SEQUENCES: 16

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/65,673

; FILING DATE: 16-JUN-1987

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 630,557

; FILING DATE: 19-JUL-1984

;SEQ ID NO:9:

; LENGTH: 357

5405942-9

Query Match 49.1%; Score 253.6; DB 6; Length 357;  
Best Local Similarity 79.2%; Pred. No. 8.8e-69;  
Matches 206; Conservative 50; Mismatches 4; Indels 0; Gaps 0;

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db      43 GGACCGGAGACGCUCUGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 102

Qy      61 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db      103 AGGGGCUUUUAUUUCAACAAGCCCACAGGGUAUGGCUCCAGCAGUCGGAGGGCGCCUCAG 162

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db      163 ACAGGUAUCGUGGAUGAGUGCUGUUUCCGGAGCUGUGAUCUAAGGAGGCUGGAGAUGUAU 222

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      :|||
Db      223 UGCGCACCCCUCAAGCCUGCCAAGUCAGCUCGUCUGUCCGUGCCCAGCGCCACACCGAC 282

Qy      241 ATGCCCAAGACCCAGAAGTA 260
      |:|

```





; LENGTH: 210  
5405942-11

Query Match 40.3%; Score 208.4; DB 6; Length 210;  
Best Local Similarity 99.5%; Pred. No. 6e-55;  
Matches 209; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
        |||
Db      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

Qy     61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
        |||
Db     61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120

Qy    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
        |||
Db    121 ACAGGTATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180

Qy    181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCT 210
        |||
Db    181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCT 210
```

RESULT 11

US-09-255-829-13

; Sequence 13, Application US/09255829

; Patent No. 6461617

; GENERAL INFORMATION:

; APPLICANT: Shone, Clifford Charles

; APPLICANT: Quinn, Conrad Padraig

; APPLICANT: Foster, Keith Alan

; TITLE OF INVENTION: Recombinant Toxin Fragments

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN, & FOX P.L.L.C.

; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600

; CITY: WASHINGTON

; STATE: DC

; COUNTRY: USA

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/255,829

; FILING DATE: 23-FEB-1999

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/GB97/02273

; FILING DATE: 22-AUG-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/782,893

; FILING DATE: 27-DEC-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: ESMOND, ROBERT W.

```

;   REGISTRATION NUMBER: 32,893
;   REFERENCE/DOCKET NUMBER: 1581.0130002
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 202-371-2600
;   TELEFAX: 202-371-2540
;   INFORMATION FOR SEQ ID NO: 13:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2862 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..2862
US-09-255-829-13

```

```

Query Match          40.3%; Score 208.4; DB 4; Length 2862;
Best Local Similarity 99.5%; Pred. No. 1.9e-54;
Matches 209; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2644 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 2703

Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2704 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 2763

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
        ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2764 ACAGGTATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 2823

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCT 210
        ||||||||||||||||||||||||||||||
Db      2824 TGCGCACCCCTCAAGCCTGCCAAGTCAGCT 2853

```

```

RESULT 12
5405942-15
;Patent No. 5405942
;   APPLICANT: BELL, GRAEME I.; RALL, LESLIE B.; MERRYWEATHER,
;JAMES P.
;   TITLE OF INVENTION: PREPRO INSULIN-LIKE GROWTH FACTORS
;I AND II
;   NUMBER OF SEQUENCES: 16
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/65,673
;   FILING DATE: 16-JUN-1987
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 630,557
;   FILING DATE: 19-JUL-1984
;SEQ ID NO:15:
;   LENGTH: 210
5405942-15

```

```

Query Match          40.0%; Score 206.8; DB 6; Length 210;

```



```

; REFERENCE/DOCKET NUMBER: 51875
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14..232
US-08-308-196A-1

```

```

Query Match          39.2%; Score 202.8; DB 1; Length 240;
Best Local Similarity 96.7%; Pred. No. 3.4e-53;
Matches 207; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
        |||
Db      17 GGACCGGAGACGCTCTGCGGGGCTGAGCTCGTGGATGCTCTGCAGTTCGTGTGTGGAGAC 76
        |||

Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
        |||
Db      77 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGACGGGCGCCTCAG 136
        |||

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
        |||
Db      137 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTCGAGATGTAT 196
        |||

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCT 214
        |||
Db      197 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTTGAT 230
        |||

```

# RESULT 14

PCT-US91-06452-1

```

; Sequence 1, Application PC/TUS9106452
; GENERAL INFORMATION:
; APPLICANT: Brierley, Russell A.
; APPLICANT: Davis, Geneva R.
; APPLICANT: Holtz, Gregory C.
; APPLICANT: Gleeson, Martin A.
; APPLICANT: Bradley, D. H.
; TITLE OF INVENTION: Production of Insulin-Like Growth
; TITLE OF INVENTION: Factor-1 in Methylophilic Yeast Cells
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 South LaSalle Street, Suite 900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/06452
; FILING DATE: 19910409
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/578,728
; FILING DATE: 04-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 51874
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)552-1311
; TELEFAX: (619)552-0095
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14..232
PCT-US91-06452-1

```

```

Query Match          39.2%; Score 202.8; DB 5; Length 240;
Best Local Similarity 96.7%; Pred. No. 3.4e-53;
Matches 207; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
        |||
Db      17 GGACCGGAGACGCTCTGCGGGGCTGAGCTCGTGGATGCTCTGCAGTTCGTGTGTGGAGAC 76
        |||

Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
        |||
Db      77 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGACGGGCGCCTCAG 136
        |||

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
        |||
Db      137 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTCGAGATGTAT 196
        |||

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCT 214
        |||
Db      197 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTTGAT 230

```

```

RESULT 15
US-09-029-267-13
; Sequence 13, Application US/09029267
; Patent No. 6107057
; GENERAL INFORMATION:
; APPLICANT: Crawford, Kenneth

```

```

; APPLICANT: Zaror, Isabel
; APPLICANT: Innis, Michael
; TITLE OF INVENTION: Pichia Secretary Leader for Protein
; TITLE OF INVENTION: Expression
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: United States
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/029,267
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Chung, Ling-Fong
; REGISTRATION NUMBER: 36,482
; REFERENCE/DOCKET NUMBER: 1165.100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2704
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic"
US-09-029-267-13

```

```

Query Match          39.2%; Score 202.8; DB 3; Length 390;
Best Local Similarity 96.7%; Pred. No. 4.2e-53;
Matches 207; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 60
      |||
Db      160 GGACCGGAGACGCTCTGCGGGGCTGAGCTCGTGGATGCTCTGCAGTTTCGTGTGTGGAGAC 219
      |||

Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db      220 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGACGGGCGCCTCAG 279
      |||

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db      280 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTCGAGATGTAT 339
      |||

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCT 214
      |||
Db      340 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTTGAT 373

```

Search completed: December 13, 2003, 11:44:49  
Job time : 49.8037 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 07:29:55 ; Search time 230.833 Seconds  
(without alignments)  
7443.919 Million cell updates/sec

Title: US-09-852-261-1  
Perfect score: 517  
Sequence: 1 ggaccggagacgctctgcgg.....tgaaatacacaagtaaacaat 517

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.		%					



1	517	100.0	517	9	US-09-852-261-1	Sequence 1, Appli
2	467.4	90.4	523	9	US-09-852-261-5	Sequence 5, Appli
3	377.2	73.0	471	9	US-09-852-261-13	Sequence 13, Appl
4	344.2	66.6	7260	10	US-09-919-497-24	Sequence 24, Appl
5	344.2	66.6	7260	10	US-09-880-107-3739	Sequence 3739, Ap
6	344.2	66.6	7260	13	US-09-873-319-707	Sequence 707, App
7	344.2	66.6	7260	13	US-09-960-706-1066	Sequence 1066, Ap
8	344.2	66.6	7260	15	US-10-136-639-4	Sequence 4, Appli
9	342.6	66.3	725	15	US-10-207-655-54	Sequence 54, Appl
10	325.2	62.9	539	9	US-09-852-261-3	Sequence 3, Appli
11	318.2	61.5	651	15	US-10-161-088-1	Sequence 1, Appli
12	285.4	55.2	612	13	US-10-251-661-7	Sequence 7, Appli
13	258.4	50.0	318	9	US-09-852-261-9	Sequence 9, Appli
14	247.8	47.9	487	9	US-09-852-261-11	Sequence 11, Appl
15	228	44.1	462	15	US-10-238-114-1	Sequence 1, Appli
16	210	40.6	210	13	US-09-807-742-18	Sequence 18, Appl
17	208.4	40.3	2862	13	US-10-241-596-13	Sequence 13, Appl
18	204.6	39.6	4532	10	US-09-930-377B-1	Sequence 1, Appli
19	203.6	39.4	210	10	US-09-930-377B-2	Sequence 2, Appli
20	202.8	39.2	390	15	US-10-179-046-13	Sequence 13, Appl
21	202	39.1	286	15	US-10-161-088-3	Sequence 3, Appli
22	183	35.4	516	13	US-10-029-386-5832	Sequence 5832, Ap
23	182	35.2	182	13	US-10-029-386-18231	Sequence 18231, A
24	140.2	27.1	213	15	US-10-076-816-9	Sequence 9, Appli
25	140.2	27.1	213	15	US-10-077-381-9	Sequence 9, Appli
26	123.8	23.9	621	9	US-09-921-398-40	Sequence 40, Appl
27	123.8	23.9	621	15	US-10-280-826-40	Sequence 40, Appl
28	108.6	21.0	480	9	US-09-921-398-38	Sequence 38, Appl
29	108.6	21.0	480	15	US-10-280-826-38	Sequence 38, Appl
30	101.2	19.6	210	13	US-09-807-742-19	Sequence 19, Appl
31	75.4	14.6	411	10	US-09-960-352-2082	Sequence 2082, Ap
32	72.4	14.0	854	10	US-09-954-531-989	Sequence 989, App
33	71.8	13.9	237	15	US-10-136-841-3	Sequence 3, Appli
c 34	70.6	13.7	447	9	US-09-922-217-917	Sequence 917, App
c 35	70.6	13.7	447	10	US-09-833-263-917	Sequence 917, App
c 36	70.6	13.7	447	14	US-10-025-380-917	Sequence 917, App
c 37	70.4	13.6	437	15	US-10-066-543-663	Sequence 663, App
c 38	70.4	13.6	493	15	US-10-066-543-997	Sequence 997, App
c 39	70.4	13.6	518	15	US-10-066-543-1040	Sequence 1040, Ap
c 40	70.4	13.6	536	15	US-10-066-543-428	Sequence 428, App
41	70.4	13.6	543	15	US-10-136-841-1	Sequence 1, Appli
c 42	70.4	13.6	549	15	US-10-066-543-478	Sequence 478, App
c 43	70.4	13.6	574	9	US-09-922-217-918	Sequence 918, App
c 44	70.4	13.6	574	10	US-09-833-263-918	Sequence 918, App
c 45	70.4	13.6	574	14	US-10-025-380-918	Sequence 918, App

#### ALIGNMENTS

#### RESULT 1

US-09-852-261-1

; Sequence 1, Application US/09852261

; Patent No. US20020083477A1

; GENERAL INFORMATION:

; APPLICANT: GOLDSPIK, GEOFFREY

```
; APPLICANT:  TERENGHI, GIORGIO
; TITLE OF INVENTION:  REPAIR OF NERVE DAMAGE
; FILE REFERENCE:  117-351
; CURRENT APPLICATION NUMBER:  US/09/852,261
; CURRENT FILING DATE:  2001-05-10
; PRIOR APPLICATION NUMBER:  GB 0011278.9
; PRIOR FILING DATE:  2000-05-10
; NUMBER OF SEQ ID NOS:  14
; SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH:  517
;   TYPE:  DNA
;   ORGANISM:  Homo sapiens
US-09-852-261-1
```

```
Query Match          100.0%;  Score 517;  DB 9;  Length 517;
Best Local Similarity 100.0%;  Pred. No. 2.4e-160;
Matches 517;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy      1  GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC  60
      |||
Db      1  GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC  60

Qy     61  AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG  120
      |||
Db     61  AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG  120

Qy    121  ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT  180
      |||
Db    121  ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT  180

Qy    181  TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC  240
      |||
Db    181  TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC  240

Qy    241  ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA  300
      |||
Db    241  ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA  300

Qy    301  AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG  360
      |||
Db    301  AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG  360

Qy    361  GATGTAGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTCT  420
      |||
Db    361  GATGTAGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTCT  420

Qy    421  GCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA  480
      |||
Db    421  GCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA  480

Qy    481  GATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT  517
      |||
Db    481  GATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT  517
```

RESULT 2

US-09-852-261-5  
; Sequence 5, Application US/09852261  
; Patent No. US20020083477A1  
; GENERAL INFORMATION:  
; APPLICANT: GOLDSPIK, GEOFFREY  
; APPLICANT: TEREINGHI, GIORGIO  
; TITLE OF INVENTION: REPAIR OF NERVE DAMAGE  
; FILE REFERENCE: 117-351  
; CURRENT APPLICATION NUMBER: US/09/852,261  
; CURRENT FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: GB 0011278.9  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 523  
; TYPE: DNA  
; ORGANISM: Oryctolagus cuniculus  
US-09-852-261-5

Query Match 90.4%; Score 467.4; DB 9; Length 523;  
Best Local Similarity 96.2%; Pred. No. 6.1e-144;  
Matches 501; Conservative 0; Mismatches 16; Indels 4; Gaps 2;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db      1 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

Qy     61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db     61 AGGGGCTTTTATTTCAACAAGCCACAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG 120

Qy    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy    181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      ||
Db    181 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCCAGCGCCACACCGAC 240

Qy    241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCA---G 297
      |||
Db    241 ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAAATGAAGTCTCAGAGG 300

Qy    298 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 357
      |||
Db    301 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 360

Qy    358 CAGGATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 416
      |||
Db    361 CAGGATGTAGGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 420

Qy    417 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGTATCACATTT 476
      |||
Db    421 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGTATCACATTT 480

Qy    477 CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 517
```

Db 481 CAAAGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 521

RESULT 3

US-09-852-261-13

; Sequence 13, Application US/09852261

; Patent No. US20020083477A1

; GENERAL INFORMATION:

; APPLICANT: GOLDSPINK, GEOFFREY

; APPLICANT: TERENCE, GIORGIO

; TITLE OF INVENTION: REPAIR OF NERVE DAMAGE

; FILE REFERENCE: 117-351

; CURRENT APPLICATION NUMBER: US/09/852,261

; CURRENT FILING DATE: 2001-05-10

; PRIOR APPLICATION NUMBER: GB 0011278.9

; PRIOR FILING DATE: 2000-05-10

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 13

; LENGTH: 471

; TYPE: DNA

; ORGANISM: Oryctolagus cuniculus

US-09-852-261-13

Query Match 73.0%; Score 377.2; DB 9; Length 471;

Best Local Similarity 87.8%; Pred. No. 4e-114;

Matches 455; Conservative 0; Mismatches 13; Indels 50; Gaps 2;

Qy 1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60  
|  
Db 1 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60  
  
Qy 61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120  
|  
Db 61 AGGGGCTTTTATTTCAACAAGCCACAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG 120  
  
Qy 121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180  
|  
Db 121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180  
  
Qy 181 TGCGCACCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240  
|  
Db 181 TGTGCACCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCAGCGCCACACCGAC 240  
  
Qy 241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300  
|  
Db 241 ATGCCCAAGACTCAG----- 255  
  
Qy 301 AGGAAAGGAAGTACATTTGAAGAACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360  
|  
Db 256 ----AAGGAAGTACATTTGAAGAACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 311  
  
Qy 361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419  
|  
Db 312 GATGTAGGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 371

```

Qy      420 TGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA 479
        |||
Db      372 TGCACAGTTACCTGTAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTCAA 431

Qy      480 AGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 517
        |||
Db      432 AGATGGCATTTCCTCCCAATGAAATACACAAGTAAACAT 469

```

RESULT 4

US-09-919-497-24

```

; Sequence 24, Application US/09919497
; Patent No. US20020106662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 7260
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-497-24

```

```

Query Match          66.6%; Score 344.2; DB 10; Length 7260;
Best Local Similarity 87.3%; Pred. No. 1.3e-102;
Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;

```

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
        |||
Db      311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 370

Qy      61 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
        |||
Db      371 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 430

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
        |||
Db      431 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
        |||
Db      491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 550

Qy      241 ATGCCCAAGACCCAGAAGTATCAGCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
        |||
Db      551 ATGCCCAAGACCCAG----- 565

Qy      301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
        |||
Db      566 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 621

```



```

      |||
Db      491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 550
Qy      241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db      551 ATGCCCAAGACCCAG----- 565
Qy      301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db      566 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 621
Qy      361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db      622 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 681
Qy      420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAATAAGTTTGATCACATTTTC 477
      |||
Db      682 TGCACGAGTTACCTGTAAACTTTGGAACACCTACCAAAAAATAAGTTTGATAACATTTA 741
Qy      478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||
Db      742 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 782

```

# RESULT 6

US-09-873-319-707

```

; Sequence 707, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; CURRENT FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; EARLIER FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 707
; LENGTH: 7260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 X57025
US-09-873-319-707

```

```

Query Match          66.6%; Score 344.2; DB 13; Length 7260;
Best Local Similarity 87.3%; Pred. No. 1.3e-102;
Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;

```

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 60
      |||

```

Db 311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 370  
 Qy 61 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120  
 |||||  
 Db 371 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 430  
 Qy 121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180  
 |||||  
 Db 431 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490  
 Qy 181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240  
 |||||  
 Db 491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 550  
 Qy 241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACCGAAGTCTCAGAGA 300  
 |||||  
 Db 551 ATGCCCAAGACCCAG----- 565  
 Qy 301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCGAGGAAACAAGAACTACAG 360  
 |||||  
 Db 566 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCGAGGAAACAAGAACTACAG 621  
 Qy 361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419  
 |||||  
 Db 622 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 681  
 Qy 420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTC 477  
 |||||  
 Db 682 TGCACGAGTTACCTGTAAACTTTGGAACACCTACCAAAAAAATAAGTTTGATAACATTTA 741  
 Qy 478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517  
 |||||  
 Db 742 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 782

RESULT 7

US-09-960-706-1066

; Sequence 1066, Application US/09960706

; Publication No. US20030134280A1

; GENERAL INFORMATION:

; APPLICANT: Munger, William E.

; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia Using

; TITLE OF INVENTION: Gene Expression Profiles

; FILE REFERENCE: 44921-5029-01US

; CURRENT APPLICATION NUMBER: US/09/960,706

; CURRENT FILING DATE: 2001-09-24

; PRIOR APPLICATION NUMBER: 60/223,323

; PRIOR FILING DATE: 2000-08-07

; PRIOR APPLICATION NUMBER: 09/873,319

; PRIOR FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 1124

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1066

; LENGTH: 7260

; TYPE: DNA

; ORGANISM: Homo sapiens



; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 X57025  
US-09-960-706-1066

Query Match 66.6%; Score 344.2; DB 13; Length 7260;  
Best Local Similarity 87.3%; Pred. No. 1.3e-102;  
Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db    311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 370

Qy      61 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db    371 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 430

Qy     121 ACAGGCATCGTGGATGAGTGCTTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    431 ACAGGCATCGTGGATGAGTGCTTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490

Qy     181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240
      |||
Db    491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 550

Qy     241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db    551 ATGCCCAAGACCCAG----- 565

Qy     301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db    566 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 621

Qy     361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db    622 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 681

Qy     420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTTC 477
      |||
Db    682 TGCACGAGTTACCTGTAACTTTGGAACACCTACCAAAAAAATAAGTTTGATAACATTTA 741

Qy     478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||
Db    742 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 782
```

RESULT 8

US-10-136-639-4

; Sequence 4, Application US/10136639

; Publication No. US20030072761A1

; GENERAL INFORMATION:

; APPLICANT: LeBowitz, Jonathan

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TARGETING PROTEINS ACROSS  
THE BLOOD BRAIN

; TITLE OF INVENTION: BARRIER

; FILE REFERENCE: SYM-008

; CURRENT APPLICATION NUMBER: US/10/136,639

; CURRENT FILING DATE: 2002-09-06

; PRIOR APPLICATION NUMBER: US 60/329,650  
; PRIOR FILING DATE: 2001-10-16  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 7260  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-136-639-4

Query Match 66.6%; Score 344.2; DB 15; Length 7260;  
Best Local Similarity 87.3%; Pred. No. 1.3e-102;  
Matches 455; Conservative 0; Mismatches 13; Indels 53; Gaps 5;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 60
      |||
Db    311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 370

Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db    371 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 430

Qy     121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db    431 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490

Qy     181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240
      |||
Db    491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 550

Qy     241 ATGCCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db    551 ATGCCCCAAGACCCAG----- 565

Qy     301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGACAGGAAACAAGAACTACAG 360
      |||
Db    566 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGACAGGAAACAAGAACTACAG 621

Qy     361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db    622 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 681

Qy     420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTC 477
      |||
Db    682 TGCACGAGTTACCTGT'TAAACTTTGGAACACCTACCAAAAAAATAAGTTTGATAACATTTA 741

Qy     478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||
Db    742 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 782
```

RESULT 9

US-10-207-655-54

; Sequence 54, Application US/10207655  
; Publication No. US20030118592A1  
; GENERAL INFORMATION:  
; APPLICANT: Ledbetter, Jeffrey A.

; APPLICANT: Hayden-Ledbetter, Martha S.  
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
; FILE REFERENCE: 390069.401C1  
; CURRENT APPLICATION NUMBER: US/10/207,655  
; CURRENT FILING DATE: 2002-07-25  
; NUMBER OF SEQ ID NOS: 426  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 54  
; LENGTH: 725  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-207-655-54

Query Match 66.3%; Score 342.6; DB 15; Length 725;  
Best Local Similarity 87.1%; Pred. No. 1.4e-102;  
Matches 454; Conservative 0; Mismatches 14; Indels 53; Gaps 5;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
      |||
Db     156 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 215
      |||

Qy     61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
      |||
Db     216 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 275
      |||

Qy     121 ACAGGCATCGTGGATGAGTGTCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
      |||
Db     276 ACAGGTATCGTGGATGAGTGTCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 335
      |||

Qy     181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
      |||
Db     336 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 395
      |||

Qy     241 ATGCCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
      |||
Db     396 ATGCCCCAAGACCCAG----- 410

Qy     301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
      |||
Db     411 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 466
      |||

Qy     361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTGCTC 419
      |||
Db     467 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTC 526
      |||

Qy     420 TGCAC-AGTTACCTG-TAAACATTGGAATACCGGCCAAAAAATAAGTTTGATCACATTTC 477
      |||
Db     527 TGCACGAGTTACCTGTTAACTTTGGAACACCTACCAAAAAAATAAGTTTGATAACATTTA 586
      |||

Qy     478 AAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
      |||
Db     587 AAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACAT 627
```

RESULT 10  
US-09-852-261-3  
; Sequence 3, Application US/09852261

; Patent No. US20020083477A1  
; GENERAL INFORMATION:  
; APPLICANT: GOLDSPINK, GEOFFREY  
; APPLICANT: TERENCE, GIORGIO  
; TITLE OF INVENTION: REPAIR OF NERVE DAMAGE  
; FILE REFERENCE: 117-351  
; CURRENT APPLICATION NUMBER: US/09/852,261  
; CURRENT FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: GB 0011278.9  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 539  
; TYPE: DNA  
; ORGANISM: Rattus sp.  
US-09-852-261-3

Query Match 62.9%; Score 325.2; DB 9; Length 539;  
Best Local Similarity 81.2%; Pred. No. 6.8e-97;  
Matches 429; Conservative 0; Mismatches 88; Indels 11; Gaps 4;

```
Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
        ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60

Qy     61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120

Qy    121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
        || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    121 ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy    181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
        || | | | | ||||| ||||| ||||| || || ||||| ||||| |||||
Db    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240

Qy    241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCA---G 297
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    241 ATGCCCAAGACTCAGAAGTCCCAGCCCCATCGACACACAAGAAAAGGAAGCTGCAAAGG 300

Qy    298 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 357
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360

Qy    358 CAGGATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 416
        ||| ||||| || || ||| || ||||| || || || || ||||| || |||||
Db    361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 420

Qy    417 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCA-----AAAAATAAGTTTGATC 470
        ||      | ||||| ||||| ||||| || || || ||||| || |||||
Db    421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480

Qy    471 ACATTTCAAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACAT 517
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACAT 528
```

```

US-10-161-088-1
; Sequence 1, Application US/10161088
; Publication No. US20030077761A1
; GENERAL INFORMATION:
; APPLICANT: Parrow, Vendela
; APPLICANT: Rosengren, Linda
; TITLE OF INVENTION: NEW METHODS
; FILE REFERENCE: 13425-111001
; CURRENT APPLICATION NUMBER: US/10/161,088
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: SE 0101934-8
; PRIOR FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(471)
US-10-161-088-1

```

Query Match 61.5%; Score 318.2; DB 15; Length 651;  
Best Local Similarity 81.7%; Pred. No. 1.5e-94;  
Matches 419; Conservative 0; Mismatches 83; Indels 11; Gaps 4;

[illegible]



```

Qy      241 ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACACGAAGTCTCAGAGA 300
          ||||||||||||
Db      487 ATGCCCAAGACCCAG----- 501

Qy      301 AGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 360
          ||||||||||||
Db      502 ----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTACAG 557

Qy      361 GATGTA-GAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTT 414
          |||||| |||||||| |||||||| |||| |||||||| ||||
Db      558 GATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTT 612

```

RESULT 13

US-09-852-261-9

```

; Sequence 9, Application US/09852261
; Patent No. US20020083477A1
; GENERAL INFORMATION:
; APPLICANT: GOLDSPIK, GEOFFREY
; APPLICANT: TERENGHI, GIORGIO
; TITLE OF INVENTION: REPAIR OF NERVE DAMAGE
; FILE REFERENCE: 117-351
; CURRENT APPLICATION NUMBER: US/09/852,261
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: GB 0011278.9
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-261-9

```

```

Query Match          50.0%; Score 258.4; DB 9; Length 318;
Best Local Similarity 99.6%; Pred. No. 6.6e-75;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
          ||||||||||||
Db      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

Qy      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120
          ||||||||||||
Db      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 120

Qy      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180
          ||||||||||||
Db      121 ACAGGCATCGTGGATGAGTGCTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180

Qy      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240
          ||||||||||||
Db      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGAC 240

Qy      241 ATGCCCAAGACCCAGAAGTA 260
          ||||||||
Db      241 ATGCCCAAGACCCAGAAGGA 260

```

US-09-852-261-11

US-09-852-261-11

Query Match 47.9%; Score 247.8; DB 9; Length 487;  
Best Local Similarity 74.5%; Pred. No. 2.6e-71;  
Matches 391; Conservative 0; Mismatches 77; Indels 57; Gaps 4;

[illegible]





Job time : 232.833 secs